UNIVERSIDAD AUTÓNOMA DE MADRID

ESCUELA POLITÉCNICA SUPERIOR

TRABAJO DE FIN DE MÁSTER

Theoretical and experimental study of P300 ERP in the context of Brain-computer interfaces.

Part II:

An experimental study of inter-and intra-subject variability based on EEG recordings.

Máster Universitario en Ingeniería Informática

Autor: Vanessa Salazar Palacios

Tutor: Francisco de Borja Rodríguez Ortiz
Cotutor: Vinicio Changoluisa Panchi

Grupo de Neurocomputación Biológica (GNB)
Dpto. Ingeniería Informática

Febrero, 2021

THEORETICAL AND EXPERIMENTAL STUDY OF P300 ERP IN THE CONTEXT OF BRAIN-COMPUTER INTERFACES.
PART II:
AN EXPERIMENTAL STUDY OF INTER-AND INTRA-SUBJECT VARIABILITY BASED ON EEG RECORDINGS.

Autor: Vanessa Salazar Palacios
Tutor: Francisco de Borja Rodríguez Ortiz
Cotutor: Vinicio Changoluisa Panchi

Grupo de Neurocomputación Biológica (GNB)
Dpto. Ingeniería Informática
ESCUELA POLITÉCNICA SUPERIOR
UNIVERSIDAD AUTÓNOMA DE MADRID

Febrero, 2021

Acknowledgement

I would like to thank the National Secretariat for Higher Education, Science, and Technology (SENESCYT, *as its Spanish acronym*) for providing funding for the development of my graduate studies.

To my family, I thank them for believing in me, for inspiring me to achieve my goals, but most of all for lighting my path with their prayers. I hope life will continue to give me opportunities to share my accomplishments with all of you. Most of all, I thank my mother for being my friend and always supporting my decisions.

Paco and Vini, thank you for sharing your knowledge and experience with me during the development of this project, it would not have been possible without your constant support and patience. It has been an honor to learn under the guidance of great teachers and great people like you.

To my friends and colleagues, thank you for making this a wonderful experience and for making our friendship last.

Finally, I would like to thank the person who has accompanied me on this journey, who always trusted me, who had the courage to take the journey and face the unknown, who knew how to be strong and keep going despite the adversity, thank you for giving me the opportunity to grow and learn, thanks to me.

Agradecimiento

Quiero empezar agradeciendo a la Secretaría Nacional de Educación Superior, Ciencia, Tecnología e Innovación (SENESCYT) por brindar financiamiento para el desarrollo de mis estudios de posgrado.

A mi familia, les agradezco por creer en mí, por inspirarme a lograr mis metas, pero sobre todo por iluminar mi camino con sus oraciones. Espero que la vida continúe dándome la oportunidad de compartir mis logros con todos ustedes. Sobre todo, agradezco a mi madre por ser mi amiga y siempre apoyar mis decisiones.

Paco y Vini, gracias por compartir conmigo sus conocimientos y experiencia durante el desarrollo de este proyecto, no hubiera sido posible sin su constante apoyo y paciencia. Ha sido un honor aprender bajo la guía de grandes maestros y grandes personas como ustedes.

A mis amigos y compañeros, gracias por hacer de esta una experiencia maravillosa, que nuestra amistad perdure.

Finalmente, quisiera agradecer a la persona que me acompañó en este viaje, quien siempre confio en mi, quien tuvo el coraje de emprender el camino y enfrentar lo desconocido, quien supo ser fuerte y seguir adelante a pesar de la adversidad, gracias por darme la oportunidad de crecer y aprender tanto, gracias a mí.

Abstract

Abstract —

The P300 event-related potential (ERP) is closely related to cognitive processes such as: attention, working memory, consciousness, among others. This signal has been shown to have a large variability that occurs independently of the subject or the cognitive process being assessed. Detecting and characterizing this variability is important for understanding cognitive changes and using this knowledge to improve the performance of P300-based brain-machine interfaces (BCIs).

There are several studies that have shown that the cognitive processes associated with P300 signal generation are highly variable between subjects, as this variability depends on factors such as: attention level, memory, age, experimental characteristics, and so on, but also that variability exists within the same subject when the same task is performed in different time periods.

For this project, a study and analysis of inter- and intra-subject variability is proposed by characterizing the P300 signal according to the specificities of each subject. For this purpose, an experimental study is proposed to be performed in the laboratory of the Biological Neurocomputation Group (GNB) of Autonomous University of Madrid on 12 subjects, using the oddball paradigm to generate visual P300 ERP by electroencephalography (EEG).

The proposed experiment was designed to evaluate variability in different circumstances: i) variability between the different subjects, ii) variability within the same subject during the 3 proposed experimental days, and iii) variability determined by the difference between the two helmets considered for the experiment due to their design and repositioning on the different experimental days. Two EEG helmets were used for this study: Enobio 8 and g-Tec g.USBAmp with dry electrodes.

For the analysis, the characterization of the signals was first performed, and the cosine distance was calculated using the coefficient of determination (r2) results for each subject to show the difference between the P300 and non-P300 signals. In addition, a selection of electrodes using the Bayesian Linear Discriminant Analysis with an exhaustive search of electrodes by ”forward selection” (BLDA-FS) is proposed to determine how the selected electrodes vary for each subject or experimental day.

After characterizing the behavior observed in the different subjects, it can be concluded that variability is widespread and undeniable both between subjects and within the same subject when performing the same task on different days. This variability is highly related to the underlying neural activity of each individual as well as to the experimental characteristics.

Moreover, the study of intra-subject variability has allowed us to observe how the results differ between the two helmets used for the experiment, highlighting that the difference between the accuracy values in the detection of P300 calculated for both helmets in the analyses performed is not representative.

The exploratory variability analysis performed allowed the attainment of a configuration that reached accuracy values up to 76.10 % for the Enobio helmet and 79.34 % for the g-Tec helmet, both with passive and active dry electrodes. It was also found that not accounting for variability can degrade the performance of a BCI system.

Key words — Brain-Computer Interface, Enobio 8, g-Tec g.USBAmp, dry electrodes, electroencephalogram, Oddball paradigm, event-related potential, P300, variability, coefficient of determina-

tion in P300, Bayesian Linear Discriminant Analysis, exhaustive search using ”forward selection” ”, electrode selection.

Resumen

Resumen —

El potencial relacionado con eventos (ERP) P300 está estrechamente relacionado con procesos cognitivos como: atención, memoria de trabajo, conciencia, etc. Se ha demostrado que esta señal tiene una gran variabilidad que se presenta independientemente del sujeto o del proceso cognitivo que se evalúe. Detectar y caracterizar esta variabilidad es importante para comprender los cambios cognitivos y utilizar este conocimiento para mejorar el rendimiento de las interfaces cerebro-máquina (BCI) basadas en P300.

Existen varios estudios han demostrado que los procesos cognitivos asociados a la generación de la señal P300 son muy variables entre sujetos, ya que esta variabilidad depende de factores como: nivel de atención, memoria, edad, características experimentales, etc., pero también que la variabilidad esta presente dentro el mismo sujeto, cuando la misma tarea se lleva a cabo en diferentes períodos de tiempo.

Para este proyecto, se propone un estudio y análisis de la variabilidad inter e intra-sujeto mediante la caracterización de la señal P300 según las especificidades de cada individuo. Para ello, se propone un estudio experimental, realizado en el laboratorio del Grupo de Neurocomputación Biológica (GNB) de la Universidad Autónoma de Madrid, sobre 12 sujetos utilizando el paradigma Oddball para la generación de ERP P300 visuales mediante electroencefalografía (EEG).

El experimento propuesto fue diseñado para evaluar la variabilidad en diferentes circunstancias: i) la variabilidad entre los diferentes sujetos, ii) la variabilidad dentro del mismo sujeto durante los 3 días de experimentación propuestos, y iii) la variabilidad determinada por la diferencia entre los dos cascos considerados para el experimento, debido a su diseño y al reposicionamiento de estos en los diferentes días de experimentación. Para este estudio se utilizaron dos cascos EEG: Enobio 8 y g-Tec g.USBAmp con electrodos secos.

Para el análisis, la caracterización de las señales se realizó en primera instancia, mediante la distancia del coseno basada en los resultados del Coeficiente de Determinación (r2) calculado para cada sujeto para mostrar la diferencia entre las señales P300 y no P300. Además, se propone una selección de electrodos utilizando el Análisis Discriminante Lineal Bayesiano mediante una búsqueda exhaustiva de electrodos a través de ”forward selection” (BLDA-FS) para determinar cómo varían los electrodos seleccionados para cada sujeto o día de experimentación.

Después de caracterizar el comportamiento observado en los diferentes sujetos, se puede concluir que la variabilidad esta ampliamente presente y es innegable, tanto entre sujetos, como dentro de un mismo sujeto al realizar la misma tarea en diferentes días. Esta variabilidad esta altamente relacionada con la actividad neuronal subyacente de cada individuo, así como también, con las características experimentales.

Además, el estudio de la variabilidad intra sujeto ha permitido observar cómo difieren los resultados entre los dos cascos utilizados para el experimento, destacando que la diferencia entre los valores de precisión en la detección de P300 calculados para ambos cascos en los análisis realizados, no es representativo.

El análisis exploratorio de variabilidad llevado a cabo permitió obtener una configuración que alcanzó valores de precisión de hasta 76,10 % para el casco Enobio y 79,34 % para el casco g-Tec,

ambos de electrodos secos pasivos y activos respectivamente, además se ha podido observar que no considerar la variabilidad puede deteriorar el rendimiento de un sistema BCI.

Palabras clave — Interfaz Cerebro-Computador, Enobio 8, g-Tec g.USBAmp, electrodos secos, electroencefalograma, paradigma oddball, potencial relacionado con eventos, P300, variabilidad, coeficiente de determinación en P300, Análisis Discriminante Lineal Bayesiano, búsqueda exhaustiva mediante ”forward selection”, selección de electrodos.

Contents

List of Figures

List of Tables

1

Introduction

1.1 Motivation

A Brain Computer Interface (BCI) is based on neural activity used as a non-muscular channel for the interaction of a subject with its environment (Allison et al., 2007, Guger et al., 2001, Wolpaw et al., 2002). In many cases, event-related potentials (ERPs) have been used as control signals for BCIs, with the P300 signal considered the most relevant for this application (Donchin et al., 2000, Sellers and Donchin, 2006).

The signal P300 event-related potential (ERP) is elicited by the oddball paradigm (Donchin and Coles, 1988), which consists of detecting the presence of an infrequent stimulus from a group of stimuli, generating attentional processes that control the perception of the stimulus change.

Cognitive processes associated with P300 signal generation are highly variable. Several studies have shown that this variability is influenced by factors, among which the following can be mentioned: attentional level (Polich, 1996), memory (Donchin, 1981), experimental characteristics (C. Guo et al., 2006, Duncan-Johnson and Donchin, 1977, Isreal et al., 1980), Inter-Stimuli Interval (Fitzgerald and Picton, 1981), electrode location (Alexander et al., 1996), pathological conditions (Sellers and Donchin, 2006), age (Houlihan et al., 1998, Emmerson et al., 1989), Polich, 1996), gender differences (Vaquero et al., 2004), etc.

In (Dixon et al., 2004) a variability classification is suggested, for this study, the inter-and intra-subject variability will be considered. According to (Hale et al., 1988), inter-subject variability refers to the differences between individuals measured in a single occasion or a single task, while intra-subject variability (S. Li et al., 2001) is defined as the differences found in measuring a single person, in a single task, on multiple occasions.

In several studies (Hertzog et al., 1992, Gordon and Carson, 1990), it has been shown that large intra-subject variability occurs in cases where some kind of underlying neural pathology is present. A comparative analysis between healthy subjects and subjects with dementia was conducted in (Hultsch et al., 2000), the results suggest that intra-subject variability remained relatively stable over time in healthy subjects, but in subjects with a pathology intra-subject variability was much higher, suggesting

that this variability is an indicator of impaired neurological mechanisms.

Regarding inter-subject variability, it is shown in (Reinhart et al., 2011, Ou et al., 2012, Polich, 1997), that P300 signal have high inter-subject variability due to age, pathological conditions, gender differences, electrode location, etc.

To reduce inter-subject variability, several techniques have been proposed, such as specific classification models (S. Lu et al., 2008), a classifier calibration strategy (Xu et al., 2016), characterization of the evolution of voltage deflections for electrode selection (Changoluisa et al., 2020,Changoluisa et al., 2018), and so on.

These results suggest that studies of inter- and intra-subject variability should be considered when analyzing BCI. Their correct interpretation may be the key to suggesting strategies that improve the performance of BCI systems, and may also lead to identifying neural pathologies in individuals and understanding underlying cognitive processes in brain function. Therefore, this project proposes to investigate and characterize this inter-and intra-subject variability in ERP generation, focusing on the P300 signal.

1.2 Aim of study

The present project proposed as a Master's thesis consists of a theoretical and experimental investigation of P300 Event Related Potentials (ERPs) in the context of brain-computer interfaces. The project is divided into two parts, the present document corresponds to the second part and focuses on an experimental investigation of inter-and intra-subject variability based on electroencephalogram (EEG) recordings.

The aim of this study is to obtain the data to be analyzed in the first phase of the experiment, which will later serve as the object of study. After obtaining the data, it is proposed to analyze the difference, in terms of accuracy, between a general analysis and a specific analysis of behavior between subjects and within the same subject. To achieve this, the following secondary objectives are proposed:

- Analyze the characteristics of the experiment to be conducted in detail, in addition to the considerations necessary for proper data acquisition.
- Examine the data obtained after preprocessing to determine the general characteristics for each subject in terms of precision.
- Propose strategies for analyzing the data obtained, moving from generality to specificity, i.e., from an analysis of inter-subject variability to an intra-subject analysis.
- Establish metrics to characterize the differences and similarities found in the analyzes.

1.3 Memory structure

This memory comprises five chapters distributed as follows:

- The introduction described in Chapter 1 contains the motivation for the proposed project and its objectives set.

- Chapter 2 of this document contains a detailed description of the variability concepts, signal acquisition, and other topics on ERPs and oddball paradigms that were not mentioned in Part 1 (Salazar, 2020). The concepts discussed in this chapter are considered necessary to understand the study conducted in this document.
- The experimental setup, signal processing, and intended variability study are described in detail in Chapter 3.
- Chapter 4 presents the results obtained from the experiments and the variability analysis for each approach considered.
- Finally, Chapter 5 describes the conclusion and discussion arising from the proposed project.
- At the end of the document are appendices that providing detailed information on the analyzes performed.

2

State of Art

This chapter describes the concepts deemed necessary to better understand the proposed project. It starts with an overview of variability, as it is the main topic on which the research focuses. First, it is mentioned in general terms and then related to the P300 signal, since it is the signal of interest.

Next, a brief introduction to the P300 signal will be given to explain the oddball paradigm in detail as this is the paradigm that will be used in the experimental part. Finally, the signal acquisition devices are considered focusing on EEG, as this is the device used for signal acquisition in this project.

2.1 Variability Analysis

It is essential to recognize the omnipresence of variability to understand it (Cobb and Moore, 1997), variability exists and is present in everything and everywhere. This recognition leads to the search for the possible sources that define this variation in the context of the study. To this end, this section briefly defines variability, focusing on ERPs.

Variability analysis can be defined as a comprehensive assessment of the degree and character of patterns of variation over time intervals (Bravi et al., 2011). This analysis found applications in many different research fields (Ho et al., 2008, Aranzana et al., 2002, Tarvainen et al., 2009, Siegler, 2007).

According to (Siegler, 2007) variability in relation to cognition exists not only at the neural and associative level, but also at the level of higher-level rules, strategies, and cognitive theories, i.e., variability exists and is present independently of the cognitive process or subject, but also within a single subject solving the same problem at close intervals. Recognizing this variability is important to accurately understand cognitive change (Damon et al., 2006).

Several studies have shown that ERPs allow the assessment of human cognitive processes with high temporal resolution (Donchin et al., 2000, Donchin and Coles, 1988). Experiments conducted to assess the behavior of ERPs using EEG are usually performed on multiple subjects and in multiple sessions. This information is generally processed using the averages method, which consists in averaging the different sessions and analyzing them in groups with the same experimental conditions,

from which the extraction of certain features that allow some generalization of the model is desired.

However, a general analysis without considering the characteristics of the different sessions loses information that might be relevant for describing the underlying cognitive process in different subjects or even within the same subject in different time periods. One of the main components of ERPs that allows a clearer interpretation of processes associated with cognition, such as attention, is the P300. This is the signal that will be used for this study.

2.1.1 Variability and P300

One way to obtain P300 ERPs is through the well-known oddballl paradigm. This paradigm consists in detecting the presence of a rare stimulus from a group of stimuli, i.e., the P300 component reflects, by positive or negative voltage deflections, an underlying brain activity induced by the presence of a stimulus. This requires the subject to make a comparison between the stimulus to be discriminated, i.e., the ”target” stimulus, and the other stimuli, thereby generating attentional processes that control the perception of the stimulus change. In (Polich, 1996) it is shown that the amplitude of the P300 signal is directly influenced by the level of attention during task development.

It is hypothesized that the properties of the P300 signal, such as amplitude or latency, depend strongly on the individual characteristics of each subject as well as on the characteristics of the experiment (Polich et al., 1990, Polich, 1996, O'Donnell et al., 1987, O'Donnell et al., 1990, J. Lu et al., 2013, Intriligator and Polich, 1995). The cognitive processes involved in the generation of a P300 signal are highly variable not only between subjects but also within a subject.

These differences between subjects become a limitation in terms of generalization. For this reason, variability is an essential parameter to consider in order to improve the characteristics of a BCI system, whose goal is adaptability in addition to high accuracy. There are several proposed classifications of the existing types of variability (Alwin, 1994, Stuss et al., 1994, Nesselroade and Featherman, 1997), but three types are proposed as the most relevant (Dixon et al., 2004) :

- Differences between individuals measured on a single occasion or task are referred to as inter-individual or inter-subject variability (Hale et al., 1988), also called diversity.
- The second type proposed is intra-individual or intra-subject differences or dispersion. This is a consideration after evaluating the variability in the performance of a single subject, only once, in many tasks (Christensen et al., 1999).
- The third type of variability is defined with the condition of measuring a single person, in a single task, in multiple occasions. The variability that exists between occasions is called intra-individual or intra-subject variability (S. Li et al., 2001).

For this study, we will focus on inter-subject variability and intra-subject variability. It is important to consider the variability caused by the repositioning of the helmets on the different experimental days, which is analyzed in the intra-subject variability type. Despite careful preparation of the experiment, it is very likely that the electrodes will not be in exactly the same positions when the helmet is repositioned as they were last time. It must also be considered that there is variability between helmets as the construction, material, and design varies from brand to brand.

2.1.2 Intra-subject variability

In neuroscience, it is common to study patients with neurological disorders who have not only low performance scores but also inconsistencies between measurements. In (Hertzog et al., 1992), it is hypothesized that large intra-subject variability in a subject indicates fluctuations between individuals mental states. In contrast, in (Gordon and Carson, 1990), high levels of intra-subject variability are found in Alzheimer's disease. A comparative analysis between healthy subjects and subjects with dementia was conducted in (Hultsch et al., 2000), the results suggest that in healthy subjects, intra-subject variability remained relatively stable over time, but in subjects with a pathology the levels of intra-subject variability were much higher, suggesting that this measure is an indicator of impaired neurological mechanisms.

In (Ouyang et al., 2017a), intra-subject variability in P300 ERPs is analyzed in a single trial, noting that ERP dataset processing typically uses large sets of trials that are expected to exhibit similar behavior, generally trials are averaged to analyze these data, but this procedure results in ignoring the variability that exists between the brain activity associated with each trial. For this reason, a trial-by-trial analysis is performed to show the differences between trials.

In (Blankertz et al., 2011) significant differences in cognitive processes between trials were shown, furthermore in (Verleger, 1997) the differences between trials regarding latency were analyzed. Studies that are considered representative have been described, but there are many more studies that related to this type of variability. These results suggest that the measure of intra-subject variability should not be ignored when analyzing a BCI system. Its correct interpretation may be the key to suggest improvements for BCI systems or, on the other hand, to identify pathologies in individuals.

2.1.3 Inter-subject variability

In (Reinhart et al., 2011, Ou et al., 2012, Polich, 2007) it is shown that the P300 wave has high inter-subject variability. Subject-related characteristics may be a cause of inter-subject variability, including age, pathological conditions, sex differences, electrode locations, etc.

One of the cases that has been studied to characterize inter-subject variability are neural networks that control cognitive processes. Because of the underlying structures, it is important to remember that they are not static, these structures vary and are reconfigured based on various factors such as: Intelligence (Schultz and Cole, 2016), cognitive activities like learning (Bassett et al., 2011), resting state (R. Zhang et al., 2015), depending on the task set (Cole et al., 2014), etc.

In (F. Li et al., 2020), a strategy defining multiple regions of interest (ROIs) is proposed to analyze this variability between activated and deactivated neural networks in each subject compared to a cognitive process. Several techniques are proposed to reduce inter-subject variability. In (S. Lu et al., 2008), a specific classification model is proposed that requires a training phase to adapt the system to each subject. In a second instance, the inclusion of a new subject in the evaluation is achieved only by learning over a subset of the measured data. The results show that this approach minimizes the inter-subject variability.

On the other hand, considering the existing variability in the amplitude and latency of the elicited signals in each electrode, as shown in (Ouyang et al., 2017b, Polich, 2007), a methodology is proposed in (Changoluisa et al., 2020), to characterize the evolution of the voltage deflections, this characterization can be applied to the selection of electrodes, as shown in the study, which in turn can be

considered as a selection of features. The proposed methodology not only improves the accuracy of the classification, but also reduces the variability between subjects, maximizing the adaptability of the BCI system. Moreover, the proposed methodology proves to be simple and reduces computational cost. In a similar study (Changoluisa et al., 2018), based on the analysis of cumulative peak difference is used as an intrinsic ERP feature for electrode selection. This proposed method contributes to the management of variability, both inter- and intra-subject, in addition to improving the accuracy of classification.

Finally, in (Xu et al., 2016), a classifier calibration strategy is proposed using inter- and intra-subject information to improve classification accuracy. Following the preceding information, it is important to recognize the importance of measuring inter-subject variability in the context of BCI. Knowing its value can help not only to implement strategies that improve BCI performance, but also to understand the underlying cognitive phenomena and brain function.

2.2 The P300-ERP

As mentioned earlier, the P300 signal is a positive deflection that can be detected in EEG recordings. It occurs with a latency of about 300 ms after the occurrence of a stimulus under certain experimental conditions, this latency value can vary in a range of 250-750 ms, depending on the individual characteristics of the subject and the experiment (Comerchero and Polich, 1999).

In several studies, the oddball paradigm is used by researchers for experimental tasks because good amplitude and latency values can be obtained with the P300 signal (Alexander et al., 1995, Furdea et al., 2009, van Dinteren et al., 2014). The specific experimental conditions necessary to obtain a P300 signal from the odd ball paradigm are described in detail below (Donchin and Coles, 1988).

2.2.1 The Oddball Paradigm

As described in Part 1 of this project (Salazar, 2020), the P300 ERP signal is a positive deflection generated after the detection of a stimulus is noticed. Typically, this signal is elicited by the oddball paradigm, which consists of presenting a target stimulus within a group of non-target stimuli. When the subject identifies this target stimulus, a P300 signal is generated. This is only possible if the subject is actively engaged in the assigned recognition task (Picton, 1992).

Studies have shown that the amplitude of a P300 signal increases with the improbability of the target stimulus and with the relevance of that stimulus relative to non-target stimuli (Wickens et al., 1983,Isreal et al., 1980). In (Duncan-Johnson and Donchin, 1977), it was shown that the amplitude of a P300 is inversely proportional to the probability of the target stimulus; this can be achieved by specific experimental conditions that make the task difficult or fast enough to redirect the subject's attention to the desired task of identifying the target stimulus.

When we speak of probability, we can refer to a temporal probability or a stimulus probability, i.e., how likely is it in the first case that a target stimulus will occur in a given period of time, or how likely is it in the second case that the stimulus will occur after a given number of non-target stimuli. In (Fitzgerald and Picton, 1981) it is suggested that the P300 is more related to temporal probability, for which the consideration of Inter-Stimulus Interval (ISI) is an elementary factor in experimental design.

With these considerations, we know that the goal of the experiment to elicit P300 ERPs is to create the necessary conditions to maximize the difference between target and non-target signals. With this idea, several types of experiments have been proposed. For this analysis, we will divide them into visual paradigms and non-visual paradigms.

2.2.1.1 Visual P300 Paradigm

In (Fazel-Rezai et al., 2012), a summary of visual paradigms is presented, and those considered most relevant are discussed in more detail below.

Row/Column

This is the most common experimental paradigm for eliciting a P300. It was presented by (Farwell and Donchin, 1988), as Farwell and Donchin speller BCI. The Row/Column System (RC) consists of a 6x6 dimensional matrix containing letters and numbers on a computer screen, each row/column is flashed randomly, the subject must fixate on one of the symbols shown and silently count how many times that stimulus appears. A P300 is elicited each time the subject locates the selected symbol, i.e., an implicit target stimulus, among the non-target stimuli. This experiment is a reference for studies on BCI systems, due to the positive results and the fact that it does not require extensive training.

In any case, there are authors who point out that this system is more efficient when a fixed stimulus or target is established, since the gaze has a lot to do with the performance of the system (Brunner et al., 2010, Treder and Blankertz, 2010). The investigation was conducted with the aim of improving the system to achieve better results it was improved and modified. Some of these newly proposed systems are described below.

Single Character

This system, unlike the traditional system, RCs flash independent characters with longer delays than RC. In the Single Character system (SC), the duration of a character is 60 ms with an interval of 40 ms between images, twice as slow as RCs, but larger P300 amplitudes could be detected than in the previous case (Guger et al., 2009).

Checkerboard

The design of the checkerboard system (CB) consists of an 8x9 matrix composed of black and white squares, with the symbols of each matrix flashing randomly and the two colors merging. The CB system aims to eliminate the repetition of elements displayed one after the other and, moreover, to reduce distraction, that is, images that can confuse the user and provoke a reaction as if they were target symbols. This is possible because this system separates the rows and columns of the matrix, in addition to the fact that it is not allowed to perform simultaneous adjacent flashes. It was shown in (Townsend et al., 2010), that this system has improvements in precision compared to the RC paradigm.

Region-Based

The Region-Based system (RB) proposed by (Fazel-Rezai et al., 2012) eliminates the idea of rows and columns previously considered, and instead flashes multiple blinking regions with grouped symbols. Character recognition is done at two stages. In the first, groups are formed and are located in different regions of the screen, the subject is asked to select a symbol and must identify it within the 7 groups that appear, if the subject can identify it a P300 signal is generated.

In the second stage, the characters of a selected group are distributed in different regions of the screen, each symbol flashes and the subject must identify the selected symbol (Fazel-Rezai et al., 2011). In (Fazel-Rezai et al., 2011), it is shown that the RB system reduces the effect of "near-target" error due to the distribution of symbols on the screen.

Alternative Visual paradigms

So far, systems similar the RC and RB systems have been analyzed, with certain variations and the RB system, which proposes a new approach based on regions. In this section, we analyze new but less common systems used to elicit a P300 ERP, also based on the oddball paradigm.

One of the approaches that stands out is the one proposed by (F. Guo et al., 2008). The proposed system has five possible targets, namely left, right, up, down and enter commands of a virtual keyboard. This system replaces the flashes with motion. Under each of the five stimuli, there is a vertical bar that moves at random intervals. In (Hong et al., 2009), a system based on matrices and flashes is compared with the moving bar system proposed by (F. Guo et al., 2008), showing that the moving stimuli cause stronger early components (N200) than the flashing stimuli.

In (Piccione et al., 2006) a new visual stimulus is proposed consisting of four arrows (up, down, left and right) randomly presented along the screen, subjects were asked to look only at the arrow indicating the direction of a moving object. This proposal required a short training period to achieve good performance. It was possible to observe through the results that the performance in subjects with some sort of pathology was lower than that of healthy subjects.

Finally, in (Hoffmann et al., 2008) a similar paradigm for eliciting P300 is proposed, consisting of six choices of images shown in random order, with an ISI of 400 ms. For this experiment, participants with a certain degree of disability and subjects without known pathologies were used. After detailed experimental development, a precision of 100% was achieved in the detection of P300 signals was achieved in 4 of the 5 disabled subjects.

The results show the success achieved with the experimental design, compared to other visual paradigms. This can be attributed to the number of images presented to the participants, after comparing it with the study conducted in (Picton, 1992) and (Sellers and Donchin, 2006) in the visual paradigm, and considering the theory of (Duncan-Johnson and Donchin, 1977) explained previously, according to which a larger amplitude is attributed to the P300 signal when the probability of occurrence of the target stimulus is low. In addition, compared to the previously explained P300 speller, replacing the area of the matrix with the multiple options presented by a single image is thought to improve subjects' ability to focus on the stimuli and improve their concentration on a single item, which is beneficial for participants with visual impairments.

2.2.1.2 Non-Visual P300 Paradigm

Despite the successes shown in visual paradigms (Brunner et al., 2010, Treder and Blankertz, 2010), performance levels have been shown to be largely dependent on vision (Kaufmann et al., 2013) suggesting that many potential users of BCI systems who have visual impairments or limitations in eye-muscle control may have difficulty performing these types of experiments. For this reason, BCI systems that do not depend on vision have been proposed.

Tactile modality

Based on the idea that tactile stimuli can successfully elicit an ERP (Ito et al., 1992, Satomi et al., 1995), a tactile system is presented in (Brouwer and Van Erp, 2010) using a vest with multiple vibration sources called tractors. Throughout the experiment, several vibration patterns were presented, with one particular pattern as the target, and the other patterns as distractors. With this study, it was found that it is feasible to implement a BCI system based on a tactile paradigm.

In (Serrano-Marugán et al., 2014), a tactile system is also presented that, unlike the previous case, presents the stimulus for the hands and not for the thorax. This experiment was conducted with a group of blind children, presenting two stimuli consisting of horizontal and vertical lines. Considering the horizontal line as the target stimulus, which was presented with a probability of 20% throughout the experiment, a sufficient probability to elicit a P300 (Johnson and Donchin, 1978). The results showed the feasibility of using a tactile system to generate P300 signals in subjects with visual impairment.

In addition to the cases described, there have been other studies based on BCI systems using a tactile paradigm, with results as successful as those using a visual stimulus (Cao et al., 2018, Guger et al., 2017, Yamaguchi and Knight, 1991, Cincotti et al., 2007), noting that systems with tactile stimulation require a longer training period.

Auditory modality

As in the previous case, due to the need to adapt BCI systems to users with visual impairments and because, as shown in (Hillyard et al., 1973, Halder et al., 2010), the potentials evoked by auditory stimuli have relevant properties, some BCI systems based on auditory stimuli are proposed.

For auditory stimuli, two modalities are considered: the binary modality, where subjects can choose between only two responses (Hill et al., n.d.), and a muti-class approach, where there are more than two stimuli (Furdea et al., 2009, Klobassa et al., 2009).

One of the first studies conducted is that of (Sellers and Donchin, 2006), this experiment consisted of a stimulus paradigm with four options ('YES', 'NO', 'PASS', 'END'), the subject was asked to consider the stimuli ' YES and 'NO' in response to a posed question, which had a probability of 25%. Although it was not compared with visual methods, it was found that this paradigm could be successfully applied to the target users, which in this case consisted of subjects with amyotrophic lateral sclerosis (ALS) and subjects without known pathologies.

Other similar auditory systems elicit an ERP based on discrimination of a target stimulus from "pips", this type of experiment is used for N1 or P2 signals (Schwent et al., 1976, Honjoh and Okita, 1981, Okita, 1979).The study conducted in (Furdea et al., 2009, Baykara et al., 2016, Klobassa et al., 2009) is an adaptation of the speller paradigm, to auditory tasks for P300 ERPs. In (Kaufmann et al., 2013) a comparative study between visual, tactile, and auditory modality systems is proposed.

2.2.2 Signal acquisition devices

As described in Part 1 of this project, signal acquisition can be performed by a variety of techniques, which are classified into three groups depending on the degree of invasiveness: invasive acquisition techniques, partially invasive acquisition techniques, and non-invasive acquisition techniques. In this case, we will focus on non-invasive techniques.

2.2.2.1 Non-Invasive acquisition techniques

This technique uses sensors that are located on the scalp. In (Nam et al., 2014), two methods of acquisition of non-invasive signals can be distinguished: i) Direct measurements, these record the electrical (EEG) or magnetic activity of the brain (MEG), and ii) Indirect measurements, which reflect the metabolism of the brain (e.g. fMRI, fNIRS and PET). For purpose of the present project, only direct measurements are described.

Magnetoencephalography

The magnetoencephalography technique (MEG) measures the magnetic fields generated by neural activity of the brain due to the small currents that occur when groups of neurons are activated due to a stimulus or spontaneous activity (Makella and Jousmaki, 2001).

Electroencephalography

EEG is a technique for monitoring electrical brain activity and can be analyzed in the time domain or in the frequency domain. In the first case, an EEG is measured as variations in voltage values at specific times in response to a stimulus. If the voltage changes after a certain time after the occurrence of an event or stimulus, it is called event-related potentials (ERPs). On the other hand, EEGs in the frequency domain are measured as voltage fluctuations at specific frequencies (Graimann et al., 2010).

As explained in (Peng et al., 2015), to record EEG signals, it is necessary to use a helmet with at least three electrodes corresponding to ground, reference, and recording electrodes. The electrodes can be made of different materials such as silver, silver chloride or gold and these can be wet or dry. In the first case, a conductive gel can be placed between the electrode and the scalp and in the second case, it is applied directly on the skin.

The use of helmets with dry electrodes offers several advantages, including the possibility of developing non-invasive BCI for people with severe disabilities by avoiding the application of gel, which reduces the preparation time before experiments and allows robust and affordable brain measurements for neuroscientific experiments as described in (Popescu et al., 2007).

Despite the more recent use of dry electrodes, research using this technology was proposed as early as the 1990s (Matsuo et al., 1973), with the aim of reducing the size of EEG devices and making their use simpler and more versatile, and of course because they are able to measure high quality signals despite hair, without possible effects on the subject's health (Grozea et al., 2011).

3

Methods: EEG helmets, our experiment design and data analysis

This chapter describes the methods used to develop the project. Given the experimental part of the proposed project, the methods are part of the development of the investigation. First, the experimental features are described detailing the experimental setup, the subjects participating in the experimental phase, and the experimental design. Then, the pre-processing of the recorded data is explained in detail. After pre-processing, an initial exploratory analysis of the data is also performed and finally the characteristics of the variability analysis conducted are described.

3.1 Characteristics of the experiment

The experimental study to be conducted in the present work was proposed in (Changoluisa, n.d.), it consists of a six-choice paradigm similar to the one proposed in (Hoffmann et al., 2008). The design of the experiment was proposed based on the objectives described in Chapter 1, that is, with the aim of performing multiple analyzes of inter-subject variability, with 12 subjects participating in the experiment, performing the same task under the same conditions.

On the other hand, the objective is to analyze the variability that can occur in the same subject performing the same activity in different time periods, in this case in 3 experimental days with a maximum interval of 2 days between experiments. Moreover, each day consists of two sessions since it is necessary to have significant data for statistical analysis. Another variability parameter that we try to analyze is the variability caused by the use of two distinct helmets due to their differences. The two helmets described below have been used for the experimentation:

- Enobio : This is a medical grade wireless system for high precision EEG recording with up to 32 channels as Figure 3.1 show. Enobio offers handy gel, solid gel and dry electrode solutions. In this case 8 dry electrodes arranged according to the international 10/20 system (Sharbrough et al., 1991) are used. Electrode selection was based on the findings of (Krusienski et al., 2006),

Figure 3.1: Enobio helmet. (Modified from (BrainSecrets, n.d.).)

which suggest that the selected electrodes gave the best overall performance, Figure 3.4a, shows their location.

- g-Tec: This is a wearable EEG headset for recording brain activity for clinical use. It consists of an electrode network with up to 32 electrodes, an amplifier (g.USBAmp), and a preamplifier (g.Sahara). In this study, 16 dry electrodes are used, their location is shown in Figure 3.2, which were also located according to the international 10-20 system, this is shown in Figure 3.4. The selection of the electrodes was based on (Guger et al., 2018, Vidal, 1977, Donchin et al., 2000).

Figure 3.2: g-Tec helmet. (Modified from (HFTCo., n.d.))

3.1.1 Experimental Setup

To develop the experiment, the subject was placed in front of a computer screen on which a group of 6 images were displayed individually, including: a monitor, a telephone, a lamp, a window, a door and a stereo; this is shown in Figure 3.3.

Figure 3.3: Set of six images used to evoke a P300.

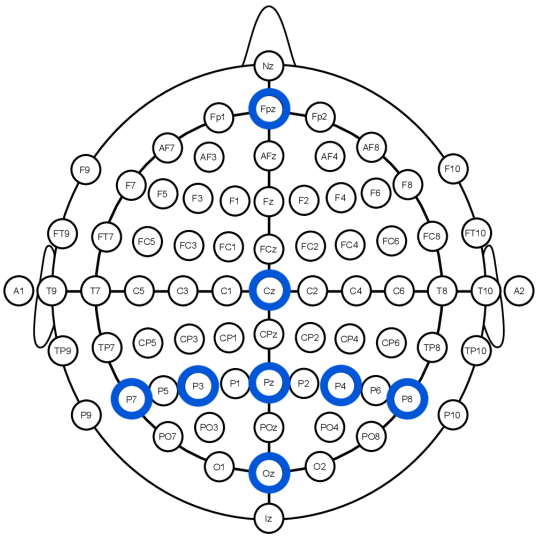
.

These images were displayed in random sequences, and the duration of the flash of each image was 100 ms, the interval between each stimulus had a duration of 300 ms, no image was shown in this interval, and the total duration of each stimulus (ISI) is 400 ms. Data acquisition was performed using the BCI2000 system (Schalk et al., 2004) and EEG signals were processed using the MATLAB software toolbox Fieldtrip.

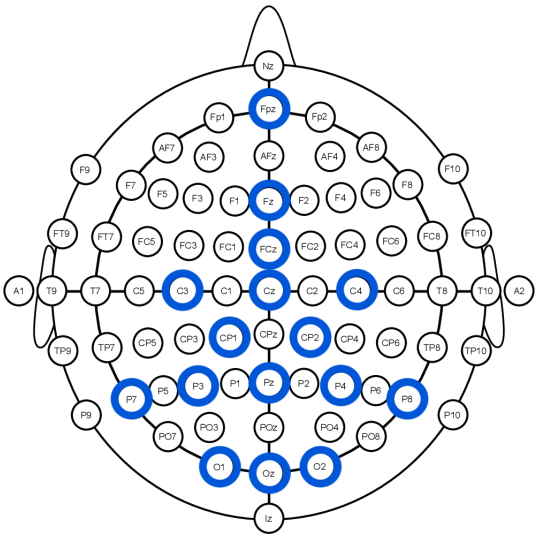
3.1.2 Subjects

The experiment was tested on a population of 12 subjects, 4 of the 12 subjects were tested for the development of this project. The data from the remaining 8 subjects are existing data from the proposed PhD thesis in (Changoluisa, n.d.).

The age of the participants ranged from 23 to 35 years (M:27.1,SD:3.57) with 16.66% identified as female and 83.33% identified as male. None of them showed any known pathology of cognitive impairment.



(a) Electrodes location in Enobio helmet, according to the international 10/20 system.



(b) Electrodes location in g-Tec helmet, according to the international 10/20 system.

Figure 3.4: Electrodes location in Enobio and g-Tec helmets.

3.1.3 Experimental Schedule

Six sessions were performed by each subject distributed over 3 days, i.e., two sessions per day. All subjects had a maximum interval of two days between sessions. For the performance of each session, the protocol explained below was taken into account before verifying that the subject had not suffered any epileptic seizures in order to perform the experiment.

- It was necessary to obtain the measurements of the head of each subject to determine the size and position of the helmet for each specific case. To conduct this experiment, the electrode placement of the international 10/20 system is used, as explained above. The goal of using these standard electrode placements is to define proportional distances between craneal landmarks for the correct positioning of the helmet in each specific case. In the case of the standard 10/20 system, proportional distances of 20 % of the total size between nasion and inion are used.

- A demo was used to explain the experiment to the subject, which was presented after the helmet and noise reduction devices were put on.

- Subjects were placed in front of a computer screen that was at the same height for all subjects, with a distance of 80 cm from the center of the screen to the subject's forehead.

- The subject was asked to count the "target" images that were presented throughout the experiment.

- The experiment consisted of the presentation of the target image at the beginning of the run, a run consisted of a group of six images presented randomly in blocks, during 15 repetitions. At the end of the run, the target image was randomly changed, and a new run was started, repeating the procedure until each of the six images had been targeted. A session consists of 6 runs, and a run consists of 15 target (P300) trials, and 75 non-target (non-P300) trials .

- At the end of each session, the subject was asked how many targets were observed during the session to evaluate the subject's performance.

The duration of each session was approximately 25 minutes, including electrode setup. Each session consisted of 540 trials, and each subject's total data consisted of 3240 trials.

3.2 Data Pre-processing

Before using the recorded signals in the following procedures, several pre-processing steps were performed. The pre-processing method used has been selected on the basis that it has proved successful with data of a similar nature (Hoffmann et al., 2008, Changoluisa et al., 2020). This steps are described in detail below.:

- **Referencing:** The two mastoid electrodes were used for referencing by the average method.
- **Filtering:** As described in Part I (Salazar, 2020), a forward–backward Butterworth bandpass filter was used. To apply this filter, the Fieldtrip BPFILTER function was used. Cut-off frequencies were set to 1.0 Hz and 12.0 Hz. The parameters set for this process are defined by: BPFILTORD, BPFILTTYPE, BPFILTDIR.
- **Downsampling:** The signal from the g-Tec helmet was downsampled from 2048 Hz to 32 Hz after applying the filter stage, and from 512 Hz to 32 Hz in the case of the Enobio helmet.
- **Visual Artifact elimination:** For artifact elimination, the Winsorize technique is used (Wilcox, 2012), where values below the 10th percentile were replaced by the 10th percentile, and similarly, values above the 90th percentile were replaced by the 90th percentile. This analysis is similar to that proposed in (Hoffmann et al., 2008).

After preprocessing, a structured dataset is obtained as shown in Figure 3.5. Each day of experimentation has 2 sessions of 6 runs each, which in turn consist of 15 trails consisting of 6 images, 5 non-target images and one target image, each trial having a duration of 1000 ms.

3.3 Initial exploration of recorded signals

After processing the data, a classification process is performed to preliminarily assess the quality of the measured data, for which a Bayesian Linear Discriminant Analysis (BLDA) is used through a cross-validation process. The theory associated with the BLDA algorithm is described in detail in Part 1 of this project (Salazar, 2020).

A cross-validation procedure is a technique used to evaluate the predictive performance of a model (F. Zhang, 2011). Training and validation datasets are needed to evaluate this performance. Cross-validation allows the evaluation of the results of a statistical analysis and ensures independence between the split of the two datasets mentioned, training and validation. In (Ivezić et al., 2014)cross-validation techniques are detailed, and some of them can be mentioned:

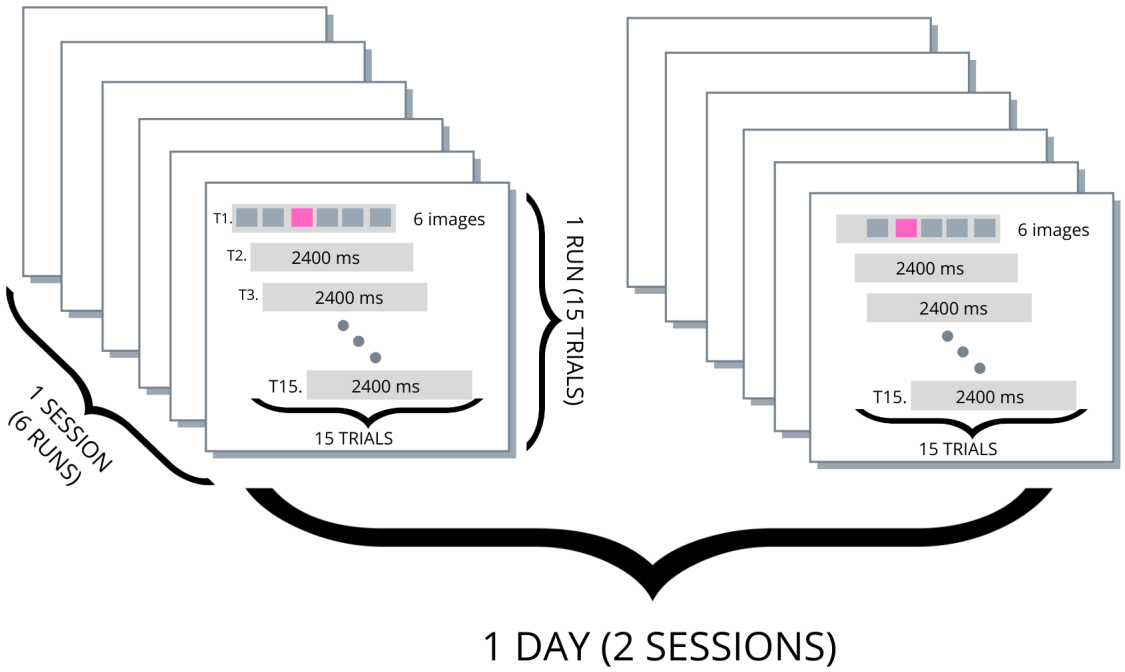


Figure 3.5: Structure of a dataset for one day of the experiment.

- Twofold cross validation.
- k-fold cross-validation
- Leave-one-out cross-validation
- Random subject cross-validation

For this project, the leave-one-out cross-validation technique is used. In this case, five sessions are used as training test and a sixth session for validation, i.e. 450 target trials and 2250 non-target trials for training; and 90 target trials and 450 non-target trials for validation. After the classifiers are trained, and applied to the validation data set, the classification process is carried out as follows:

- The validation session consists of 6 runs, each with 15 blocks of 6 images (1 target and 5 non-target).
- The single trials of each run are classified, resulting in 15 blocks of classifier outputs, each block contains 6 values (one for each run), resulting from assigning a value of '1' when the output of the classifier is equal to the target of that block and "0" otherwise.
- The values resulting from each classification run are added, and the process is repeated by iterating the validation set from the 6 possible sessions.
- When all possible variants have been considered, the accuracy is calculated after averaging the previously calculated "correct" values.

3.4 Variability Analyses

Three factors are taken into account for the variability analysis. First, a general analysis. Here the individual characteristics of the subjects are not considered in the classification process. Second, an intra-subject variability study is proposed, where the features of each subject are analyzed individually for accuracy calculation. Finally, an inter-subject study is proposed where the behavior of each subject on 3 different days is evaluated to analyze the characteristics of the subject across the different days of the experiment.

For the general analysis, all sessions of all subjects are considered, i.e., a total of 72 sessions together. For the intra-subject analysis, the 6 sessions of each subject are considered for an individual analysis per subject. For the inter-subject analysis, 3 days of the experiment are considered, each day consisting of 2 sessions: for day 1, sessions 1 and 2; for day 2, sessions 3 and 4; and for day 3, sessions 5 and 6.

Finally, using the results obtained, an attempt is also made to analyze the differences between the two helmets. In addition to the above considerations, 2 methods for electrode selection are proposed:

- The first approach consists in using the standard electrodes.
- The second approach consists in selecting the electrodes using the BLDA algorithm.

3.4.1 Standard electrodes

The selection of standard electrodes was based on exhaustive research using methods such as forward selection or backward elimination (Changoluisa et al., 2020). The results of these analyzes are described in the Table 3.1. It is observed that the results are consistent between the different authors, highlighting the presence of electrodes from both the central lobe and the occipital and parietal lobes.

Reference	Electrodes
Vidal, 1977	Pz, Oz, O1, O2, Fz
Farwell and Donchin, 1988	Pz
Polikoff et al., n.d.	Pz, Cz, Fz
Donchin and Coles, 1988	Pz, Cz, Fz, O1, O2
Kaper et al., 2004	Pz, Cz, Fz, Oz, P3, P4, PO7, PO8, C3, C4
Serby et al., 2005	Pz, Cz, Fz, Oz
Sellers and Donchin, 2006	Pz, Cz, Fz
Piccione et al., 2006	Pz, Cz, Fz, Oz
Krusienski et al., 2006	Pz, Cz, Fz, Oz, P3, P4, PO7, PO8
Hoffmann et al., 2008	Pz, Cz, Fz, Oz, P3, P4, P7, P8
Cecotti et al., 2011	Pz, Oz, P3, P7, P8
Ryan et al., 2017	Pz, Cz, Fz, Oz, P3, P4, PO7, PO8
Sheng et al., 2018	Pz, Cz, Fz, Oz, O1, O2

Table 3.1: Standard electrode selection for the detection of P300 ERPs. Approaches by different authors. Modified from (Changoluisa et al., 2020).

The data obtained from the two helmets are composed by a different number and location of electrodes, so the following electrodes were set as standard electrodes for both cases: Pz, Cz, Oz, P4, P3, P8, P7 and Fpz, as shown in Figure 3.4a. Electrode Fpz does not correspond to the usual standard electrodes chosen, but is due to the relevant activity detected in the frontal area electrodes in (Changoluisa et al., 2020). Because of the location of the electrode, it was considered to perform a preprocessing aimed at eliminating ocular artifacts using winsorizing, as explained in Chapter 3, Section 3.2. This technique is similar to that used in (Hoffmann et al., 2008) for visual artifact elimination.

Although standard electrodes without specific information about the characteristics of the experiment or the subject improve the accuracy of the results, they are not necessarily the electrodes that better capture the information to be measured in every case. For this purpose, an individual selection of electrodes is proposed that is better adapted to the individual characteristics of the subject, despite the computational cost involved.

3.4.2 BLDA and Forward Selection

In order to adapt the results to the individual characteristics of the data collected for each subject, a selection of electrodes is made using the BLDA algorithm, which makes it possible to determine, through the calculated accuracy values, which are the electrodes that maximize this accuracy and to use them for the variability analysis.

Forward selection is used for electrode selection using BLDA (BLDA-FS). Forward selection is a straightforward variable selection strategy (Herrera et al., 2009). In (Chandra, 2016), the steps followed for forward selection are described as follows:

- The most significant feature $S_1 = f_i$ is first selected based on a criterion.
- Then pairs of features with f_i are formed and the best pair is selected as $S_2 = f_i, f_j$.
- A set of 3 features with S_2 is formed and the best set of 3 features is selected as $S_3 = f_i, f_j, f_k$.
- This process is repeated until a predefined number of features are selected.

The forward selection used in this project aims to select a first electrode that maximizes the accuracy results and later combine it with a second electrode from the remaining electrodes. After measuring the accuracy results obtained with all variants, the second electrode is selected. Once the pair of electrodes is obtained, the process is repeated to obtain a third electrode and so on until the desired number of electrodes is obtained, in this case 8. The results found in the development of this project are described in detail in the next chapter.

4

Results

This chapter describes the results obtained after the analysis performed. First, an initial study of the data is conducted to determine in a general way the preliminary results of the different subjects. Then, the variability study is started, which consists of conducting a general analysis. Later an analysis of the variability between and within subjects, to end with a generalization analysis.

4.1 Initial exploration of recorded signals

Before analyzing the variability, a brief analysis of the data obtained from each subject is proposed. For this purpose, the calculation of the Coefficient of Determination (r2) is presented as a first consideration to evaluate the quality of the signal, the theory related to the r2 is explained in detail in Part 1. Figure 4.1, 4.2, and Figure 4.3 show the results of this calculation for the Enobio helmet, the g-Tec helmet with the same electrodes as Enobio, and the g-Tec helmet with 16 electrodes, respectively. For this calculation, 6 sessions with 6 runs for each of the subjects are used.

In most cases, the observed results show a clear distinction between the P300 and non-P300 signals in the range between 300 and 600 ms. This distinction is very clear in some subjects, for example: with Enobio helmet in subjects 1, 2, 6, 8 and 12, the activity is clearly distinguishable; however, in subjects 3, 4, 5, 7, 9, 10 and 11, it is more difficult to detect. In contrast, for the g-Tec helmet, activity is clearly distinguishable in subjects 1, 2, 3, 6, 8, and 12; however, it is difficult to detect in subjects 4, 5, 7, 9, 10, and 11.

To analyze how the behavior of a subject changes compared to other subjects using the same helmet and different helmets, an additional similarity analysis is proposed to analyze how similar subjects and helmets are in comparison. For this, it is proposed to use the same 8 electrodes for both helmets. Figure 4.4, shows the results of this study for which cosine distance was used. Cosine distance determines the similarity that exists between the data being compared, in this case the comparison of the data from r2 belonging to each subject. Given two vectors, A and B, the cosine distance measure can be calculated as follows:

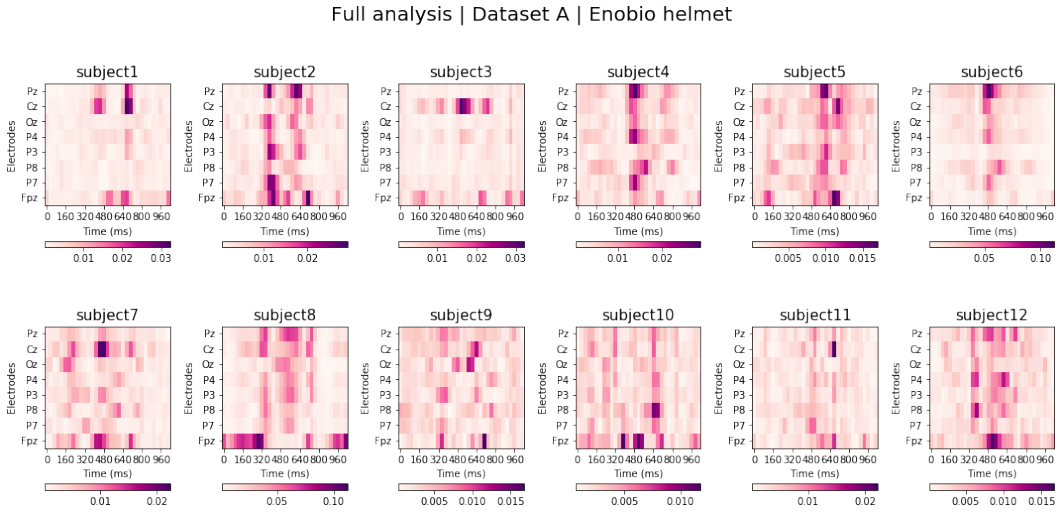


Figure 4.1: R2 values for 12 subjects with Enobio helmet, with 6 sessions per subject.

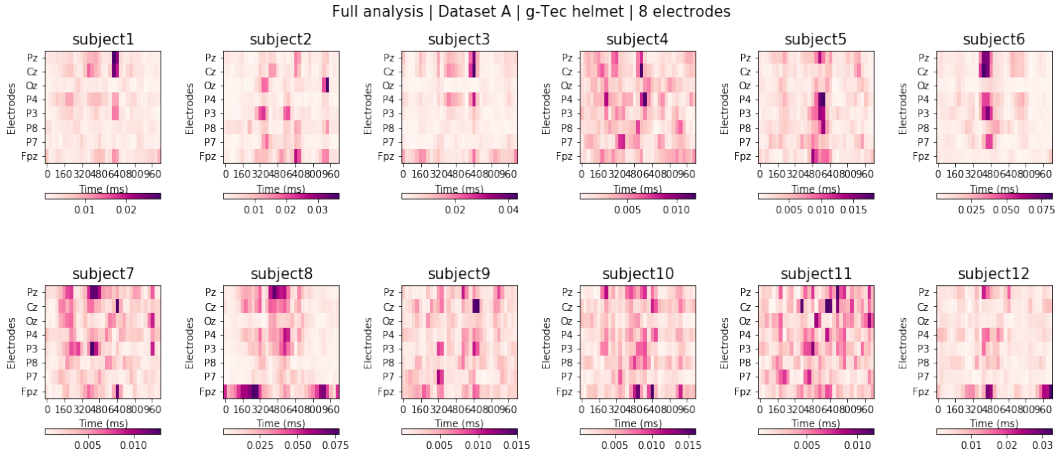


Figure 4.2: R2 values for 12 subjects with g-Tec helmet using 8 electrodes, with 6 sessions per subject.

$$cosine\ distance = 1 - \frac{A \cdot B}{||A||x||B||}. \tag{4.1}$$

Cosine distance is zero for identical vectors and 1 for orthogonal vectors (Sidorov et al., 2014). A result close to zero means that the compared data are very similar, and a result close to 1 means that the compared data are different. Before calculating the similarity values, a preprocessing of the r2 values was performed. A simple mean filter explained in González, 2008 was used considering the mean of 5 values for each point, this process is performed with the aim of smoothing the signal to reduce possible shifts.

It can be seen from the figure that in the first case, when the subjects of the Enobio helmet are compared, subjects 1, 3, 6, and 8 are the most different from the other subjects, while subjects 5 and 11 seems to be the most similar to other subjects of the same helmet. It is observed that subjects 10, 11, and 12 have a cosine distance result closer to zero with respect to the other subjects, this may be due to the fact that the data collection of these subjects was done under different conditions than the

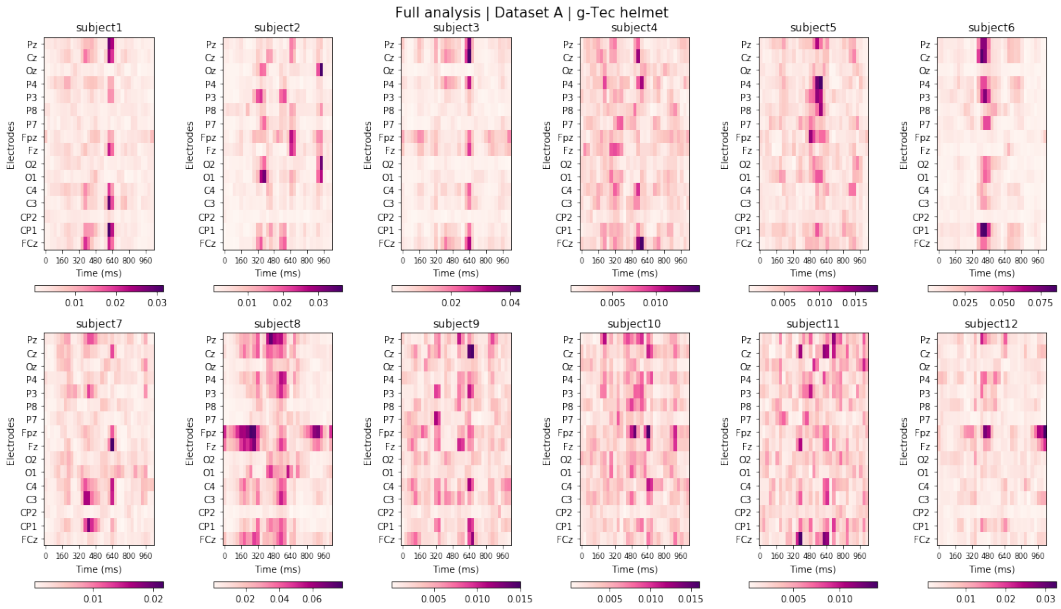


Figure 4.3: R2 values for 12 subjects with g-Tec helmet using 16 electrodes.

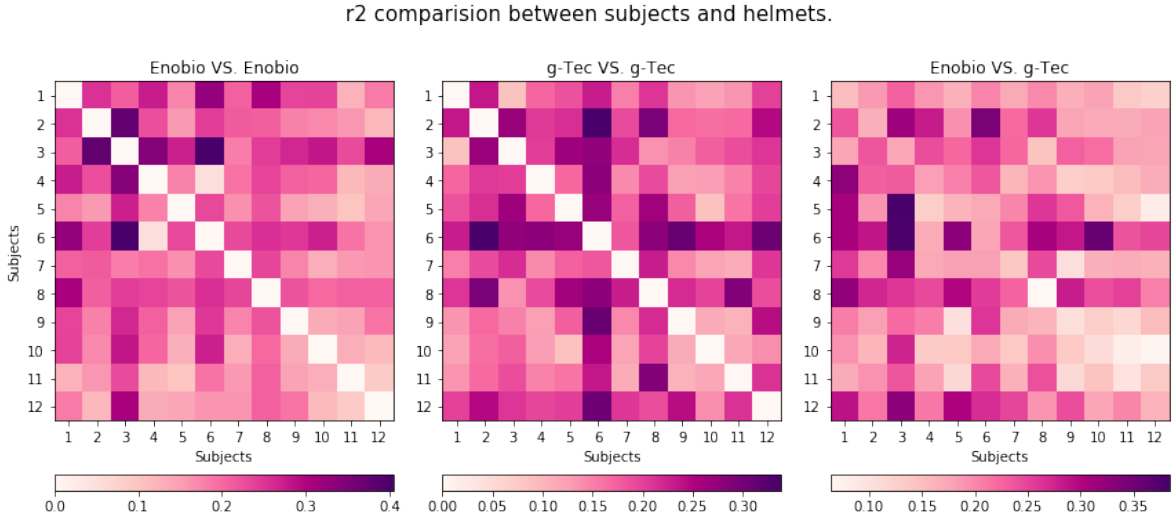


Figure 4.4: R2 comparison between subjects and helmets.

others.

In the second case, comparing the subjects of the g-Tec helmet, we can see that subjects 4, 7, 9, 10, and 11 are very similar to the other subjects, while subjects 6 and 8 seem to be different in comparison.

Finally, subjects from both helmets are compared. It is obvious that there will be a difference between the data of the subjects in one helmet and the other, due to several factors, including the repositioning of the helmet and the difference between the helmets used, such as the material and the electrodes used (there are passive electrodes in the Enobio helmet, while the electrodes on the g-Tec helmet are active). In this analysis, it is observed that the results between the same subjects on the different helmets shown on the diagonal are very similar, especially for subject 8, whose similarity

value is close to 0. It is also observed that subjects 2, 6, 8 and 12 are different from the other subjects.

In recent years, there has been increasing interest in studying the effects of individual variability in the activation and connectivity of brain regions (Demuru and Fraschini, 2020), to this end EEG signals have been used in several studies (Meir-Hasson et al., 2014, Chan et al., 2018, Palaniappan et al., 2002, Fraschini et al., 2015, Thomas and Vinod, 2016), with the aim of exploiting physiological characteristics for each subject to identify them from the rest. In (Palaniappan et al., 2002) it was shown that the levels of thought-process differ between subjects even with similar mental activity. From the r2 comparison between subjects and helmets, it was found that there are several subjects, such as subjects 3, 6, and 8, to name a few, who exhibit a characteristic behavior that is repeated in the 3 analyzes performed. This identified behavior in specific subjects could be described as a unique characteristic of each subject that could be used for recognition, like a biological fingerprint of each subject.

To complement this initial analysis, a more specific study was proposed using the BLDA algorithm as a classification strategy to measure accuracy in each case. To apply the BLDA algorithm, a cross-validation 6 is proposed using 5 of the 6 possible sessions of each subject as training and the remaining session as validation. For this classification, all available electrodes were used in each case, that is:

- Enobio helmet: Pz, Cz, Oz, P4, P3, P8, P7, Fpz.
- g-Tec helmet: Pz, Cz, Oz, P4, P3, P8, P7, Fpz, Fz, O2, O1, C4, C3, CP2, CP1, FCz.

Figure 4.5, shows the accuracy calculated for each subject. It can be observed that the results vary between helmets, and optimal accuracy is not achieved in all cases. The Table 4.1 shows the subjects ordered from highest to lowest according to the calculated accuracy values.

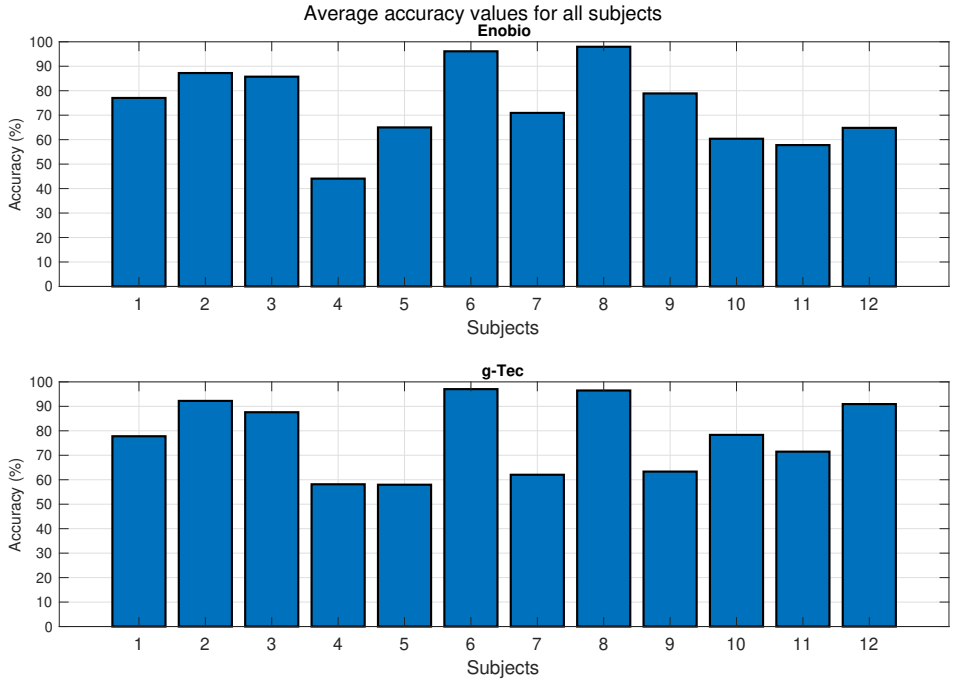


Figure 4.5: Preliminary accuracy results for subjects 1-12, considering both helmets, using 8 electrodes for Enobio helmet and 16 electrodes for g-Tec helmet.

Enobio helmet	Accuracy	g-Tec helmet	Accuracy
8	97.96%	6	97.04 %
6	96.11%	8	96.48%
2	87.22%	2	92.22%
3	85.74%	12	90.92%
9	78.88%	3	87.59%
1	77.04%	10	78.33%
7	70.92%	1	77.78%
5	65%	11	71.48%
12	64.81%	9	63.33%
10	60.37%	7	62.04%
11	57.78%	4	58.15%
4	44.07%	5	57.96%

Table 4.1: Subjects ordered according to accuracy values.

As reference, a classification accuracy of 70% is defined as a threshold value, this value has been used in several studies as an indicator of satisfactory communication for a BCI system (Furdea et al., 2009, Kübler et al., 2005, Kübler et al., 2001, Nijboer et al., 2008, Käthner et al., 2013, Perelmouter and Birbaumer, 2000).

With respect to the Enobio helmet, subject 8 shows the highest accuracy, followed by subject 6. Furthermore, subjects 5, 12, 10, 11 and 4 show accuracy values below the reference threshold. Only 7 out of 12 subjects (58%) exceed the accuracy threshold set for satisfactory communication with the Enobio helmet.

Regarding the g-Tec helmet, subject 6 obtained the best accuracy results, followed by subject 8, these two subjects show the highest accuracy values in the case of both helmets. On the other hand, subjects 9, 7, 4, and 5 are below the accuracy threshold. For the g-Tec helmet, 8 out of 12 subjects (66.6%) have a recommended accuracy level to establish satisfactory communication in a BCI system.

As can be seen in the Table 4.1, the difference between the accuracy values for the first 3 subjects in both helmets is similar, their order changes by a minimal difference. Subjects 1 and 3 also have similar accuracy values for both helmets. The other subjects show larger differences between the helmets.

Comparing this analysis with the results observed when using r2 was used and shown in Figure 4.4, we note that the subjects that show lower similarity, i.e., subjects 1, 3, 6, and 8 for the Enobio helmet and subjects 2, 3, 6, 8, and 12 for the g-Tec helmet, are the same ones that show the highest accuracy values in the analysis performed with BLDA, as shown in Figure 4.5. In the following, some specific features mentioned by the subjects during the development of the conducted experiments are detailed below.

Subject 9 stated during the second session with the g-Tec helmet that he could not identify the target image at the end of session 2. This subject stated on the second day of the experiment that he had only a few hours of sleep and that the g-Tec helmet was uncomfortable to wear. On the last day of this subject's experiment there was noise from the air conditioning unit, the subject stated that he was sweating during the test. This information could explain the low accuracy results for this subject when using the g-Tec helmet.

Subject 10, on the other hand, was visually fatigued on days 1 and 3 of the experiment. The subject wore contact lenses that forced him to blink constantly, which was a distraction in identifying the target image. In the second session of the first day, the subject counted an incorrect number of target images. On the last day of the experiment, while wearing the Enobio helmet, there was noise outside the experimental room that distracted the subject.

The count of target images for subject 11 was not correct for session 2 wearing the Enobio helmet. This subject also reported that he was unfocused during sessions 5 and 6 with the same helmet. This could explain the subject's poor performance when using the Enobio helmet. Subject 12 was unable to determine the correct number of target images during session 1 and also reported that there were sources of distraction during day 3 with the Enobio helmet.

4.2 Variability Analysis

This section contains the variability analysis performed, which is divided into 4 subsections. In a first approach, variability is analyzed in general, without considering differences between subjects or within the same subject. Then, an inter-subject analysis is performed, taking into account the variability between the different subjects participating in the experiment. Then, an intra-subject variability analysis is performed, taking into account the different days of the experiment. Finally, the circumstances under which it is possible to generalize the model by combining data from several days is analyzed.

4.2.1 General Analysis

For a first analysis of variability, a general analysis was considered to be performed in which neither inter- nor intra-subject differences are taken into account. In this first case, the standard electrodes are used to evaluate the resulting accuracy values for each subject. The selection of standard electrodes is described in Chapter 2.

The BLDA classification algorithm was used and two types of cross validation were proposed: cross validation 12 with the sessions, and cross validation 72 with the runs. Both of them correspond to the leave-one-out type of cross-validation (LOOCV). This type of cross-validation corresponds to a case of 'leave-p-out' cross-validation, in this case, p = 1, i.e. the statistic is calculated on the one 'leave-out' (omitted) sample.

Each subject has 6 sessions and each of these sessions has 6 runs, so the total number of runs for the 12 subjects is 72. Cross-validation 12 consists of iteratively using the 6 sessions of 11 subjects as training set and the 6 sessions corresponding to the one remaining subject as validation set. Figure 4.6 illustrates the results obtained. The accuracy results of Figure 4.6 are 57.93% for the Enobio case and 62.21% for the g-Tec helmet.

On the other hand, cross-validation 72 was performed, it consists of the iterative combination of 71 training runs and a validation run. The result of this analysis is shown in Figure 4.7. The accuracy values obtained of Figure 4.7 are 62.4% in the case of the Enobio helmet and 68.75% in the case of the g-Tec helmet.

In this analysis, we observed how the accuracy values vary depending on the type of cross-

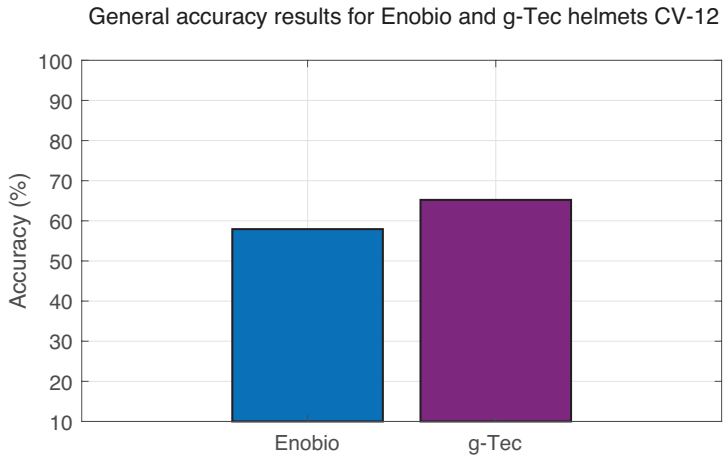


Figure 4.6: Accuracy results of general analysis for Enobio (blue) and g-Tec (purple) helmets using standard electrodes and cross-validation 12 (11 sessions used as training set and 1 session used as validation set).

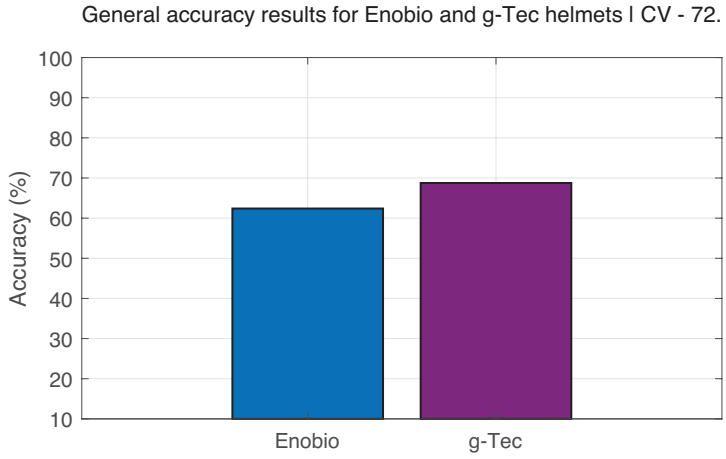


Figure 4.7: Accuracy results of general analysis for Enobio (blue) and g-Tec (purple) helmets using standard electrodes and cross-validation 72 (71 runs used as training set and 1 run used as validation set).

validation used. In the first case, when the cross-validation completely excludes all the information of one of the subjects for the classification process, the results are low according to this analysis. However, in this case, when the cross-validation includes all the information from all but one of the subjects for one run, we can see how the accuracy results, although not optimal, improve significantly compared to the previous case. When the system has more information from all users, the results are better because the system gets more information to learn.

4.2.2 Inter-subject Variability Analysis

Inter-subject analysis aims to analyze the behavior of each subject and how this behavior differs from that observed in the general analysis. Initially, an analysis is performed with standard electrodes, and later a selection of electrodes with BLDA-FS is performed, taking into account 2 types of cross-

validation.

4.2.2.1 Standard Electrodes for Inter-subject Variability Analysis

The first consideration raised was the use of standard electrodes for the classification process. As in the previous case, the BLDA algorithm was used for this process. Considering that each subject has a total of 6 sessions and each session has 6 runs, cross-validation 6 is proposed, i.e., 5 sessions are used for training and 1 for testing iteratively for each subject.

From this analysis the results obtained are shown in Figure 4.8. The upper part of the figure is the same as in Figure 4.5, but was included in this analysis to compare it with the results of the g-Tec helmet, where the set of electrodes was limited to 8 for this analysis.

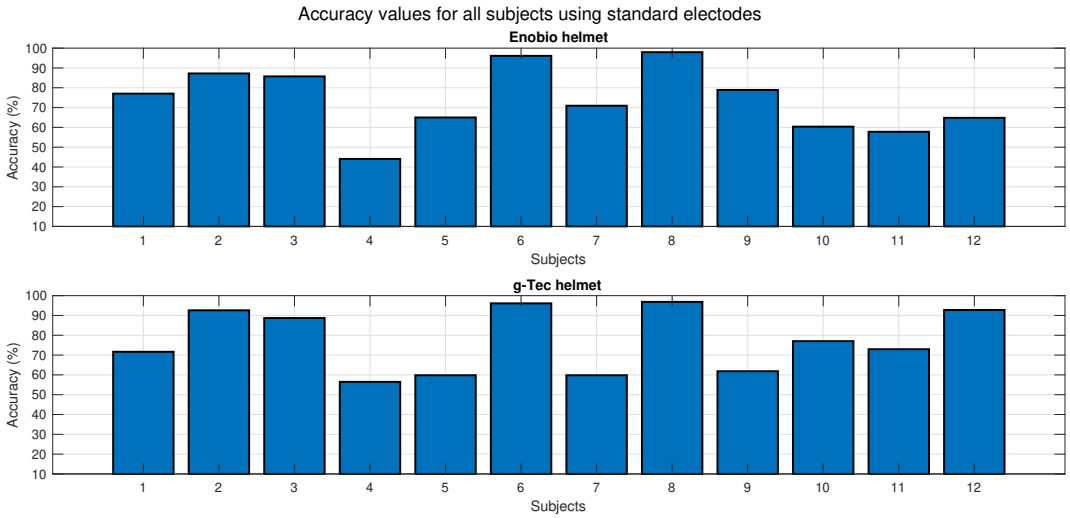


Figure 4.8: Accuracy results of inter-subject variability analysis for all subjects using standard electrodes.

The results obtained show the behavior of each subject, which improves in most cases compared to the general analysis when the individual characteristics of each subject are taken into account for the classification process. In both helmets, the subjects that preset a better accuracy are subjects 6 and 8. On the other hand, subject 4 is the one with the lowest accuracy value.

In order to compare the results of the individual helmets, an average was calculated in each case, which is shown in the Figure 4.9. From these results it can be seen that the g-Tec helmet gives a higher accuracy result, but the calculated variances rule out a significant difference.

The resulting values of the calculated average for the Enobio and the g-Tec helmet considering all subjects are 73.8% and 77.22%, respectively. It can be observed that when considering the individual characteristics of each subject in the classification process, the accuracy increases significantly compared to the general analysis by about 10% in both cases. These results show the importance of considering the individual characteristics of the subjects in order to maximize the accuracy of the results due to the large variability that exists between them.

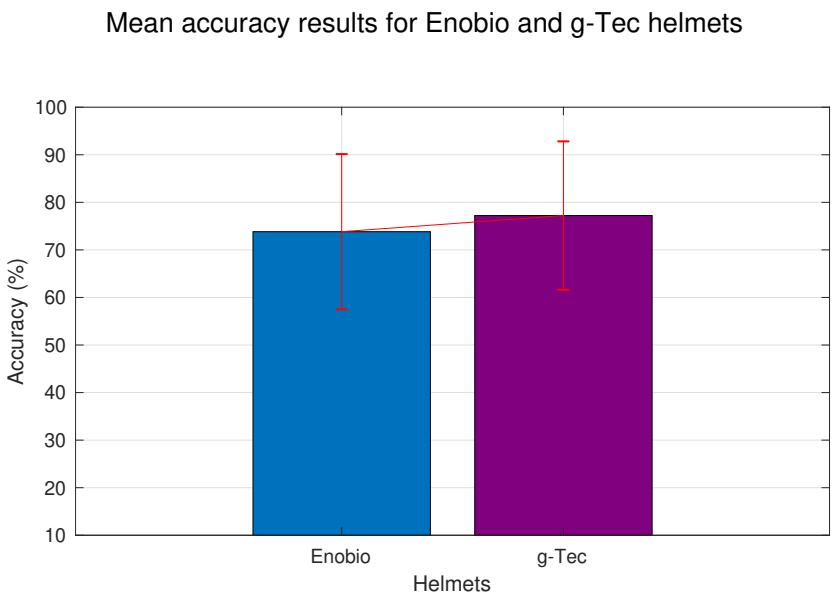


Figure 4.9: Accuracy results of inter-subject variability analysis for Enobio and g-Tec helmets using standard electrodes.

4.2.2.2 BLDA-FS Electrodes for Inter-subject Variability Analysis

On the other hand, electrode selection using BLDA-FS has been proposed to analyze inter-subject variability. The selection is done by forward selection as explained in Section 3 of Chapter 2.

The electrodes of the g-Tec helmet were limited to the 8 electrodes of the Enobio helmet, with the aim that both helmets come together under the same conditions for further evaluation. For this analysis, the use cross validation 6 is proposed, that is, 5 of the 6 sessions will be considered for training and the remaining session for validation. The accuracy values for each helmet resulting from the classification performed are shown in Figures 4.10 and Figure 4.11. Accuracy was measured by varying the number of electrodes from 1 to 8.

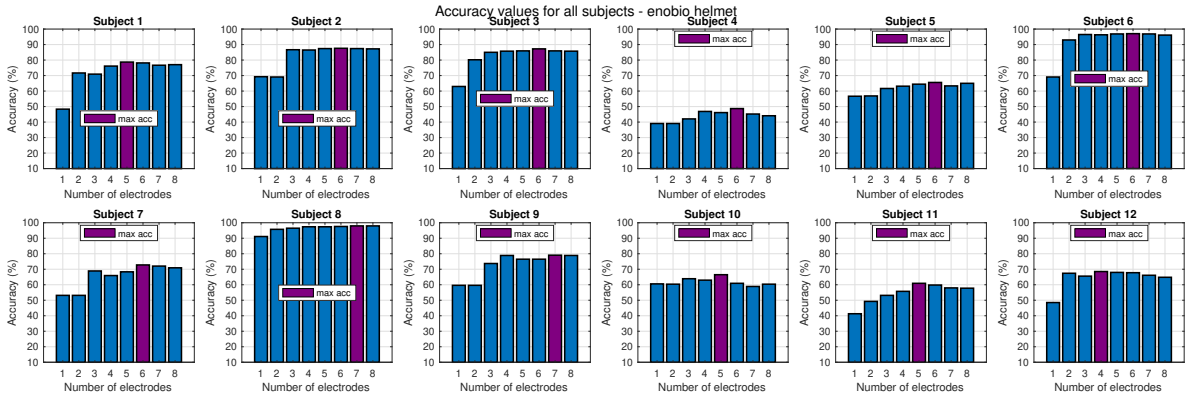


Figure 4.10: Accuracy results of inter-subject variability analysis using electrodes selected by BLDA-FS (Table 4.3) for Enobio helmet.

From this analysis, the optimal number of electrodes for each subject was defined. This is given by the number of electrodes that provide the maximum accuracy per subject. Table 4.2 shows the

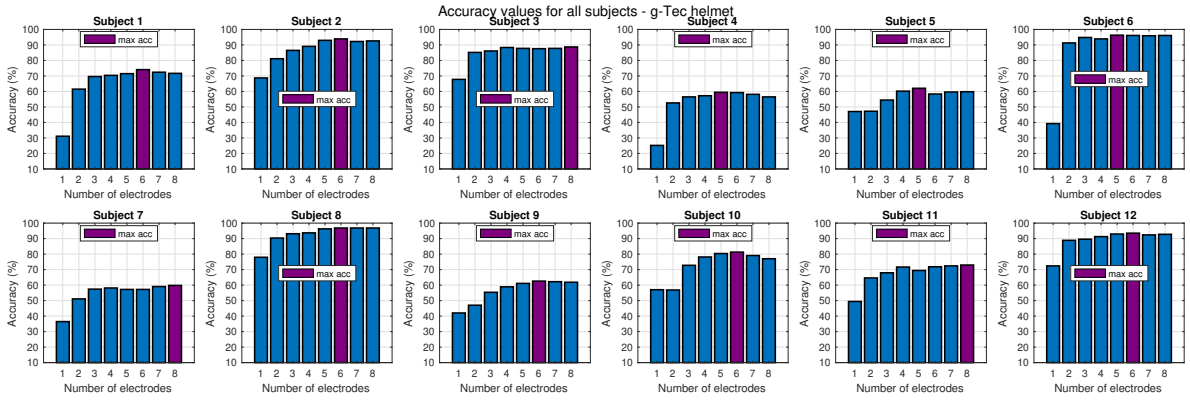


Figure 4.11: Accuracy results of inter-subject variability analysis using electrodes selected by BLDA-FS(Table 4.3) for g-Tec helmet.

optimal number of electrodes for each subject, the average for each helmet is about 6 for each case. The electrodes resulting from this analysis are shown in the Table 4.3.

Helmet	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	Mean	Variance
Enobio	5	6	6	6	6	6	6	7	7	5	5	4	5.75	0.75
g-Tec	6	6	8	5	5	5	8	6	6	6	8	6	6.25	1.295

Table 4.2: Optimal number of electrodes for each subject for the general analysis.

Subjects	Enobio Helmet	g-Tec Helmet
1	Cz, Fpz, Pz, Oz, P8, P7, P4, P3	Cz, Pz, P7, P8, Fpz, P3, Oz, P4
2	Fpz, Oz, P7, Cz, P4, P3, Pz, P8	P8, Fpz, P7, Cz, P3, P4, Pz, Oz
3	Cz, Pz, Oz, Fpz, P4, P7, P3, P8	Cz, Fpz, P4, P7, P3, Oz, Pz, P8
4	Fpz, P7, P3, P8, Cz, P4, Pz, Oz	Cz, P4, Oz, P7, Pz, P8, Fpz, P3
5	Fpz, Pz, P3, P8, P4, Cz, Oz, P7	Fpz, P3, Pz, P7, P8, Cz, Oz, P4
6	Pz, P8, Fpz, Cz, Oz, P3, P7, P4	Cz, P4, Fpz, Pz, P3, P7, P8, Oz
7	Fpz, Oz, P7, P4, Cz, P8, Pz, P3	Pz, P3, P8, Fpz, Oz, P4, P7, Cz
8	Cz, Pz, Oz, P3, P8, P4, P7, Fpz	P4, Cz, P3, P8, Pz, Fpz, Oz, P7
9	Fpz, Oz, P8, P7, Cz, P4, P3 , Pz	Cz, P8, Pz, P4, P7, P3, Oz, Fpz
10	Fpz, P7, P8, Cz, P3, Pz, Oz, P4	Fpz, P4, P8, Oz, Pz, P7, Cz, P3
11	Cz, Oz, P7, P8, Fpz, Pz, P4, P3	Cz, P4, Pz, Fpz, P8, Oz, P3, P7
12	Pz, Fpz, P4, P8, P3, Oz, P7, Cz	Pz, Fpz, P4, Oz, P8, P7, Cz, P3

Table 4.3: Electrodes selected using BLDA-FS for inter-subject analysis.

The Figure 4.12 shows the number of electrode repetitions found, corresponding to the electrodes considered optimal for each helmet. It can be observed that for Enobio helmet, the electrodes with the highest number of repetitions are Cz, Fpz and P8, while for the g-Tec helmet it is the Pz electrode.

The electrodes with the highest number of occurrences coincide with the electrodes listed in Table 3.1, especially Pz and Cz, which are considered important by several authors. In addition, in (Chang-

oluisa et al., 2020) the electrodes from the frontal area have shown great prominence, the electrode Fpz represents the electrodes in this area and it is evident in this analysis that it is also important for most subjects. Although they are not the electrodes with the maximum occurrence, it can be observed that in both cases the parietal electrodes P7 and P8 remain constant and have a great importance. In both cases, the maximum possible number of occurrences is 12, since there are 12 subjects.

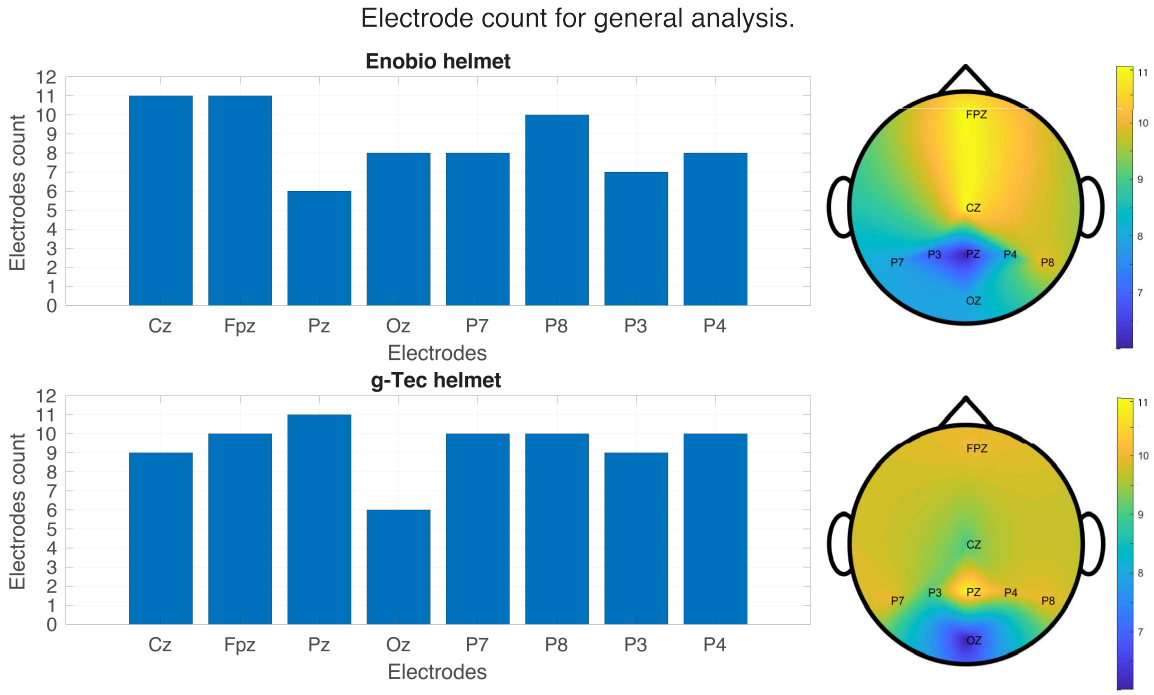


Figure 4.12: Electrode count for general analysis using BLDA-FS selection.

4.2.3 Intra-subject Variability Analysis

To analyze intra-subject variability, one analysis per day is proposed. The dataset consists of 3 experimental days, each day has two sessions, and each session has 6 runs, i.e., 12 runs per day as Figure 3.5 shows. For this analysis, standard electrodes are first used. Then, a selection of electrodes with BLDA-FS is made, for which two types of cross-validation are proposed.

4.2.3.1 Standard Electrodes for Intra-subject Variability Analysis

For this first analysis, the standard electrodes and BLDA algorithm were used for classification with cross validation 2, 1 session is used for training and 1 for testing. Figure 4.13 shows the result obtained from this analysis.

From these results, it is evident that the accuracy behavior of the different subjects varies across days. Subject 4 showed low accuracy values in relation to the other subjects in the general analysis. When analyzed daily, it can be seen that on day 1 and 2 the accuracy is very low for both helmets, but on day 3 the results improve. Subject 9 clearly has low accuracy on day 2, which could be due to

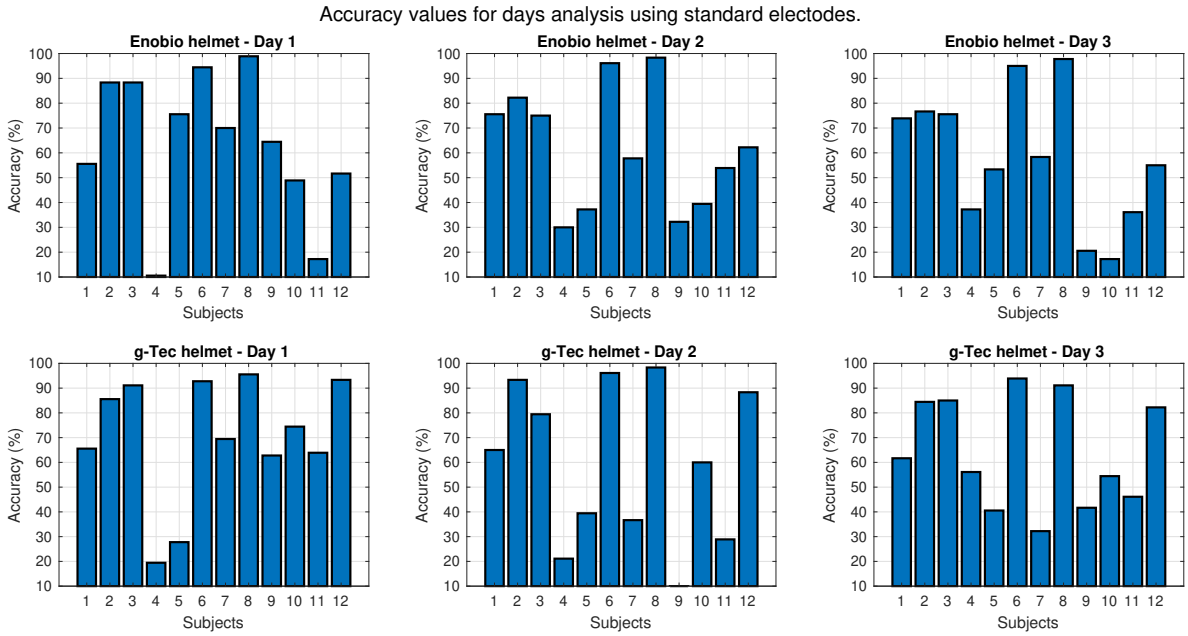


Figure 4.13: Accuracy results of intra-subject variability analysis for all subjects using standard electrodes and cross-validation 2 (1 session used as training set and 1 session used as validation set).

the fact that the subject reported being tired and not having had enough hours of sleep on the second day of the experiment. Also, on day 3, there was noise during the experiment and the subject reported that he was sweating. On the other hand, on day 3 of the experiment, noises were heard by subject 10 outside the experimental room, which, based on the results observed on that day, were clearly constituted a disturbance for the subject. Subject 11 was unable to count the correct number of target images on session 2 using the Enobio helmet, which may explain this subject's poor performance on day 1. Finally, subject 12 was unable to count the correct number of target images in session 1, which affected the first day's results.

In addition, an analysis of the behavior of the helmets is intended. For this purpose the values obtained from all subjects for each day were averaged. The results of this analysis are shown in the Figure 4.14, it can be seen that for day 1 and day 3 the g-Tec helmet presents higher accuracy values than the Enobio helmet, however the difference is not clear.

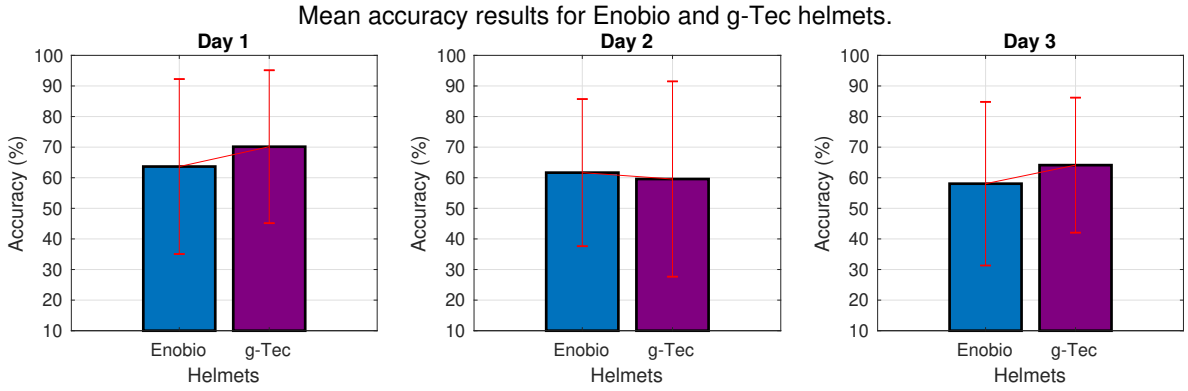


Figure 4.14: Accuracy results of inter-subject variability analysis for Enobio and g-Tec helmets using standard electrodes and cross-validation 2 (1 session used as training set and 1 session used as validation set).

A second analysis was proposed which consists in calculating the accuracy values for each day, in this case the BLDA algorithm considers a cross-validation 12, which takes into account the 12 runs of a day. The Figure 4.15 shows the results obtained from this analysis.

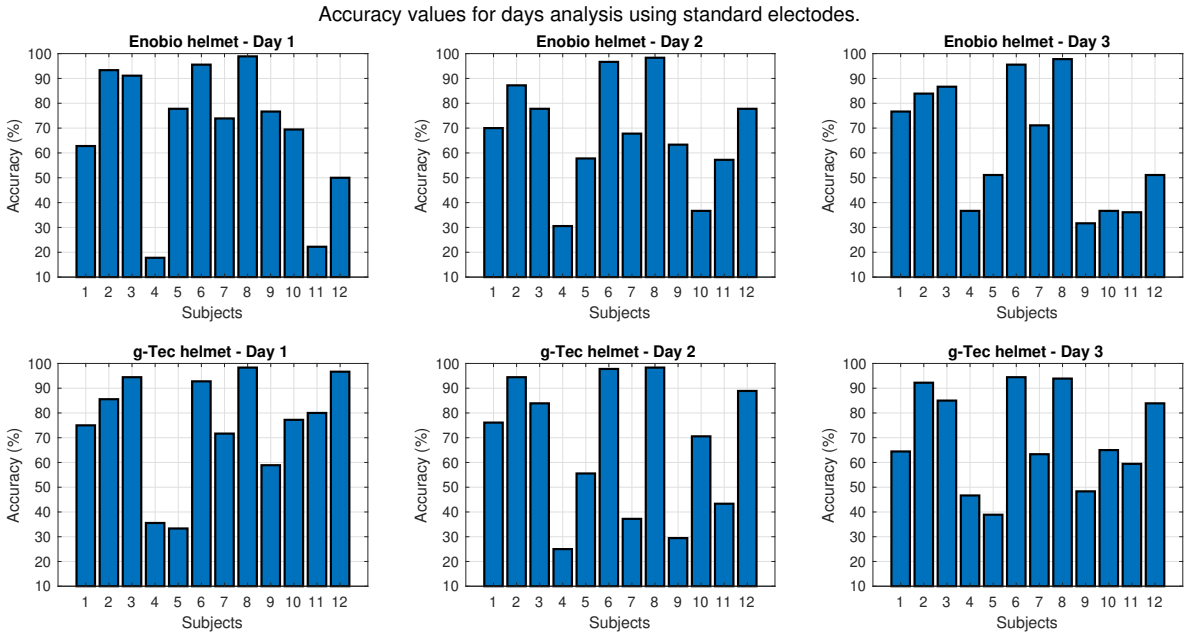


Figure 4.15: Accuracy results of intra-subject variability analysis for all subjects using standard electrodes and cross-validation 12 (11 runs used as training set and 1 run used as validation set).

With this validation, higher accuracy results are obtained for most subjects than in the previous case, while generally maintaining the behavioral pattern. For example, subject 9, although having a higher accuracy than in the above case, still has the lowest accuracy on day 2 compared to the other days, which continues to affect the average result for each helmet. The Enobio helmet has a higher accuracy for day 2, unlike the other days where the g-Tec helmet predominates, however, as in the previous case, this difference is not crucial as shown in the Figure 4.16.

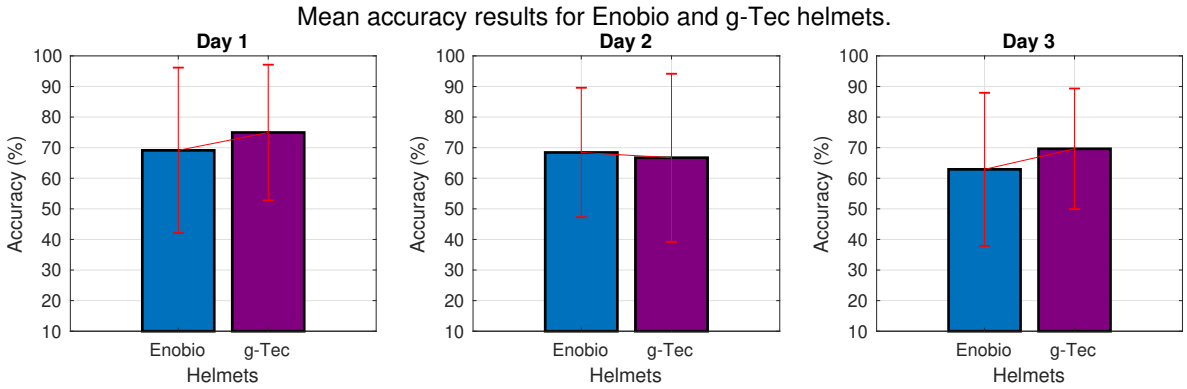


Figure 4.16: Accuracy results of intra-subject variability analysis for Enobio and g-Tec helmets using standard electrodes and cross-validation 12 (11 runs used as training set and 1 run used as validation set).

4.2.3.2 BLDA-FS electrodes for Intra-subject Variability Analysis

In addition to the analysis using standard electrodes, an analysis of the classification results measured with selected electrodes using the BLDA-FS algorithm as explained in Section 3 of Chapter 2 is proposed.

First, an analysis using cross-validation 2 is proposed, using one session for training and the rest for testing. The obtained results are shown in Appendix B, Figure B.1 and Figure B.2. A brief summary of the results obtained is presented in Figure 4.17.

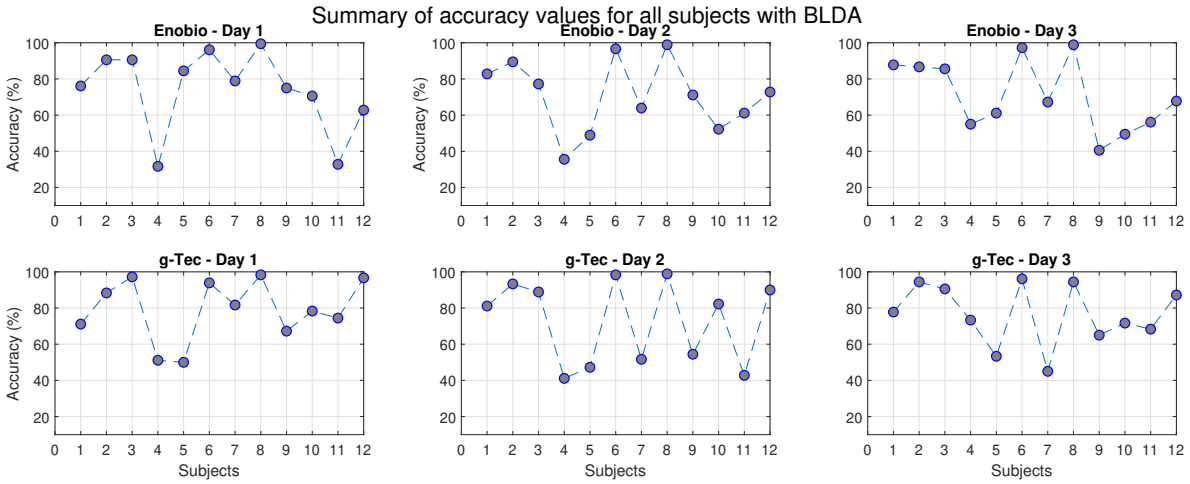


Figure 4.17: Accuracy results of intra-subject variability analysis for Enobio and g-Tec helmets with electrodes selected using BLDA-FS and cross-validation 2 (1 session used as training set and 1 session used as validation set).

In the daily analysis, as opposed to the general analysis, it is observed that fewer electrodes are required to achieve maximum accuracy. From this analysis it was possible to calculate the number of optimal electrodes, which are shown in Table 4.4. It is observed that the average number of electrodes for both helmets to maximize the accuracy value is about 4. It can be observed that for the g-Tec helmet the number of optimal electrodes considered is slightly higher. The electrodes selected from this analysis are shown in detail in Table 4.7.

Day	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	Mean	Variance
Enobio														
D1	3	3	4	2	3	5	3	2	3	2	1	5	3	1.45
D2	5	3	5	3	3	2	6	3	2	2	6	3	3.58	2.26
D3	5	4	2	4	3	4	5	4	5	2	3	4	3.75	1.11
g-Tec														
D1	4	4	3	4	4	4	4	2	4	4	5	6	4	0.90
D2	3	8	2	3	5	3	5	5	3	3	2	6	4	3.27
D3	3	4	3	4	3	5	4	4	5	4	2	6	3.92	1.17

Table 4.4: Optimal number of electrodes for each subject for intra-subject analysis using cross-validation 2 (1 session used as training set and 1 session used as validation set). This analysis is shown Appendix B in Figures B.1 and B.2.

Sbj	Enobio Helmet Electrodes Day 1	Electrodes Day 2	Electrodes Day 3
1	Pz, Fpz, P7, P8, Cz, Oz, P3, P4	Pz, Fpz, Cz, P8, P4, P3, Oz, P7	Cz, P4, P7, Fpz, P8, Pz, P3, Oz
2	Fpz, Cz, P3, P8, Pz, P4, P7, Oz	Fpz, Oz, P7, Cz, Pz, P3, P4, P8	Fpz, Oz, P7, P8, P3, Pz, Cz, P4
3	Cz, P8, Fpz, P7, Oz, P4, Pz, P3	Fpz, P4, P8, Cz, P7, P3, Pz, Oz	Cz, Fpz, P7, P3, Pz, Oz, P4, P8
4	Pz, P4, Fpz, Oz, P3, P8, Cz, P7	P7, Cz, Fpz, P8, P3, Oz, P4, Pz	P3, Oz, Fpz, P7, Cz, P8, Pz, P4
5	Fpz, Oz, P3, P4, Pz, P8, P7, Cz	P7, Fpz, Pz, Cz, Oz, P3, P8, P4	Fpz, P3, P7, Pz, Cz, P8, P4, Oz
6	Pz, P8, Fpz, P3, Cz, Oz, P7, P4	Pz, Cz, P7, Fpz, P4, P8, Oz, P3	Cz, P8, Fpz, Pz, Oz, P7, P3, P4
7	Cz, P7, Oz, P3, Fpz, Pz, P8, P4	Fpz, P4, P8, Cz, Oz, P3, P7, Pz	Fpz, P8, Cz, P7, Pz, P4, P3, Oz
8	Pz, Fpz, Cz, P3, P7, Oz, P4, P8	Pz, Cz, P4, P3, Oz, P8, Fpz, P7	P7, Pz, Cz, P3, Oz, P4, P8, Fpz
9	Fpz, P3, Oz, P8, Cz, Pz, P4, P7	Fpz, Oz, P4, P7, Cz, P8, P3, Pz	Pz, Fpz, P8, Cz, P7, Oz, P4, P3
10	Fpz, P8, Oz, Pz, P3, Cz, P4, P7	Fpz, Pz, Cz, P8, P7, P4, Oz, P3	Fpz, Oz, Pz, Cz, P4, P3, P8, P7
11	P8, Oz, Fpz, Pz, Cz, P3, P4, P7	Fpz, Oz, P7, P4, P8, P3, Cz, Pz	Fpz, P7, Cz, P3, Pz, P4, Oz, P8
12	Pz, Fpz, Oz, Cz, P7, P8, P4, P3	P8, Fpz, P3, Oz, Pz, Cz, P7, P4	Pz, Fpz, P4, P8, P3, Oz, Cz, P7
Sbj	g-Tec Helmet Electrodes Day 1	Electrodes Day 2	Electrodes Day 3
1	Cz, Pz, Fpz, P3, P4, Oz, P7, P8	Cz, P4, P3, P7, Fpz, Pz, P8, Oz	Pz, Cz, P4, Fpz, P7, P3, P8, Oz
2	Fpz, P3, Cz, P8, Pz, P7, P4, Oz	P8, Cz, Oz, Fpz, P7, Pz, P4, P3	Fpz, P3, Cz, P8, Pz, Oz, P7, P4
3	Cz, P7, Oz, Fpz, Pz, P4, P3, P8	P4, Cz, P3, Pz, Oz, P8, P7, Fpz	P4, Fpz, Pz, P7, P8, Cz, Oz, P3
4	P4, Pz, P8, Cz, Fpz, P3, Oz, P7	P4, P3, P8, Cz, Pz, Fpz, Oz, P7	P4, Cz, Fpz, P7, P8, Oz, P3, Pz
5	Fpz, P7, Pz, Oz, Cz, P3, P8, P4	Pz, P7, P4, Fpz, P3, Oz, Cz, P8	Pz, Cz, Oz, P8, Fpz, P3, P7, P4
6	P4, Cz, Pz, P3, P7, Fpz, Oz, P8	Cz, P3, P4, P8, P7, Fpz, Pz, Oz	Pz, P4, Cz, P7, P8, P3, Oz, Fpz
7	Pz, P3, P4, P8, Oz, P7, Cz, Fpz	Cz, P3, P8, P7, Oz, P4, Pz, Fpz	P3, Oz, P8, Cz, Pz, Fpz, P7, P4
8	Cz, Pz, P8, P7, Oz, P4, P3, Fpz	Cz, Fpz, Oz, Pz, P8, P4, P7, P3	Cz, P3, P7, P8, Oz, Fpz, P4, Pz
9	Pz, P3, Fpz, P4, Cz, P7, P8, Oz	Fpz, Oz, P8, P7, P4, P3, Cz, Pz	Cz, Fpz, P8, Pz, P7, P4, P3, Oz
10	Cz, Fpz, P8, P3, P4, Oz, Pz, P7	Oz, Fpz, P7, Pz, Cz, P4, P8, P3	Fpz, P3, P8, P4, Pz, Oz, P7, Cz
11	Pz, Cz, P4, Oz, P7, P3, P8, Fpz	P8, Cz, Pz, P3, Fpz, P7, P4, Oz	Cz, P7, P8, P3, Pz, P4, Fpz, Oz
12	Pz, P8, Fpz, P3, Oz, P4, Cz, P7	Cz, Fpz, P8, P4, P7, Pz, Oz, P3	Pz, Fpz, Oz, P8, P7, P3, P4, Cz

Table 4.5: Electrodes selected by BLDA-FS and cross-validation 2 (1 session used as training set and 1 session used as validation set) for intra-subject analysis.

To measure the variation of electrodes from one day to another, the optimal electrodes for each day and for each subject were considered. The Levenshtein distance is considered as a metric. The Levenshtein distance proposed in (Levenshtein, 1966), can be used to compare two strings, and its value indicates the number of deletions, insertions, or substitutions of a single character that must be made to convert one string into another. Figure 4.18 shows the results obtained from this analysis.

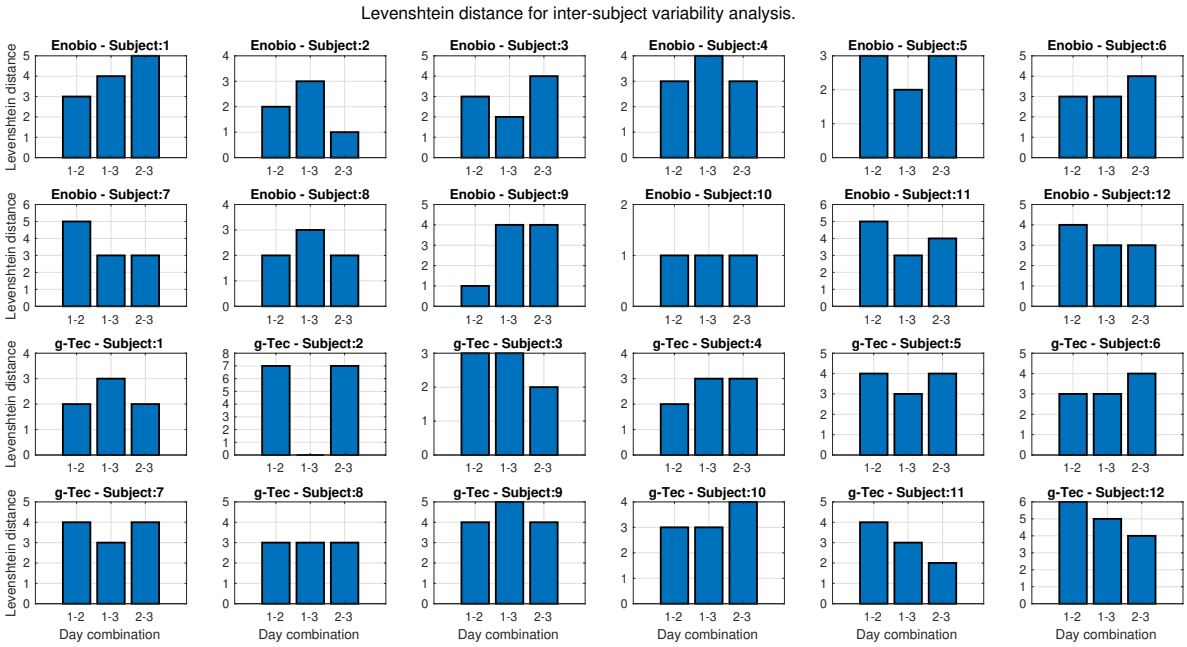


Figure 4.18: Levenshtein Distance results for intra-subject variability using cross-validation 2 (1 session used as training set and 1 session used as validation set).

As can be seen in most cases, there is a variation in the optimal electrodes selected between the different days. The only case where the electrodes are the same for two days is the case of subject 2 between days 2 and 3 in the g-Tec helmet, the other days and other subjects show differences.

A second analysis is proposed using runs for cross-validation instead of daily sessions, using 11 runs for training and 1 run for validation, iteratively varied. The results obtained from this analysis can be seen in Appendix B, Figure B.3 and Figure B.4. A summary of the results obtained is shown in Figure 4.19.

Compared to the previous results, it can be observed that the accuracy values in this proposed classification increase in most cases, it is also worth mentioning that the number of optimal electrodes in this case is lower than in the previous ones in some cases, but the mean optimal electrode mean value is still about 4. The Table 4.6 shows the selected optimal electrode values. The electrodes selected from this analysis are shown in detail in Table 4.7.

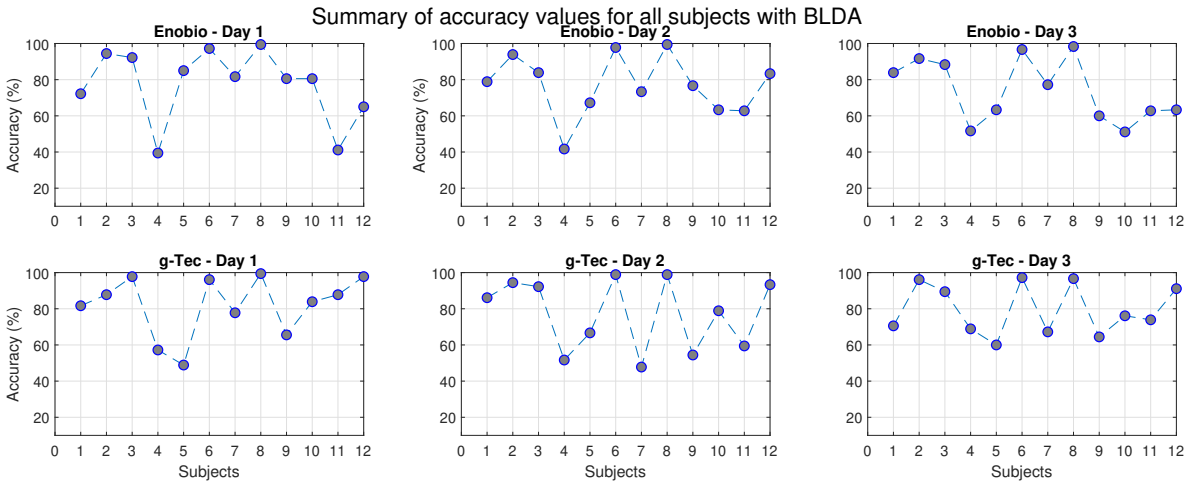


Figure 4.19: Accuracy results of inter-subject variability analysis for Enobio and g-Tec helmets using BLDA-FS electrodes and cross-validation 12 (11 runs used as training set and 1 run used as validation set).

Day	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	Mean	Variance
Enobio														
D1	2	6	4	2	3	4	4	3	5	4	1	4	3.5	1.91
D2	3	4	4	3	4	3	3	5	3	2	3	4	3.41	0.62
D3	5	4	6	2	3	5	5	6	2	1	3	3	3.75	2.75
g-Tec														
D1	5	6	2	3	3	4	3	4	5	3	4	4	3.83	1.24
D2	4	8	3	3	4	3	3	4	3	6	4	4	4.08	2.26
D3	3	5	3	3	4	4	5	2	4	5	4	3	3.75	0.93

Table 4.6: Optimal number of electrodes for each subject for inter-subject analysis using cross-validation 12 (11 runs used as training set and 1 run used as validation set).This analysis is shown Appendix B in Figures B.3 and B.4.

Sbj	Enobio Helmet Electrodes Day 1	Electrodes Day 2	Electrodes Day 3
1	Pz, Fpz, Oz, P8, P7, Cz, P4, P3	Pz, Fpz, Cz, P8, P4, P3, Oz, P7	Cz, P4, P7, Fpz, P8, Pz, P3, Oz
2	Fpz, Pz, P3, P8, P7, Cz, Oz, P4	Fpz, Oz, P7, Cz, Pz, P3, P4, P8	Fpz, Oz, P7, P8, P3, Pz, Cz, P4
3	Cz, P8, Fpz, P7, Oz, P4, Pz, P3	Fpz, P4, P8, Cz, P7, P3, Pz, Oz	Cz, Fpz, P7, P3, Pz, Oz, P4, P8
4	Pz, P4, Fpz, Oz, P3, P8, Cz, P7	P7, Cz, Fpz, P8, P3, Oz, P4, Pz	P3, Oz, Fpz, P7, Cz, P8, Pz, P4
5	Fpz, Oz, P3, P4, Pz, P8, P7, Cz	P7, Fpz, Pz, Cz, Oz, P3, P8, P4	Fpz, P3, P7, Pz, Cz, P8, P4, Oz
6	Pz, P8, Fpz, P3, Cz, Oz, P7, P4	Pz, Cz, P7, Fpz, P4, P8, Oz, P3	Cz, P8, Fpz, Pz, Oz, P7, P3, P4
7	Cz, P7, Oz, P3, Fpz, Pz, P8, P4	Fpz, P4, P8, Cz, Oz, P3, P7, Pz	Fpz, P8, Cz, P7, Pz, P4, P3, Oz
8	Pz, Fpz, Cz, P3, P7, Oz, P4, P8	Pz, Cz, P4, P3, Oz, P8, Fpz, P7	P7, Pz, Cz, P3, Oz, P4, P8, Fpz
9	Fpz, P3, Oz, P8, Cz, Pz, P4, P7	Fpz, Oz, P4, P7, Cz, P8, P3, Pz	Pz, Fpz, P8, Cz, P7, Oz, P4, P3
10	Fpz, P8, Oz, Pz, P3, Cz, P4, P7	Fpz, Pz, Cz, P8, P7, P4, Oz, P3	Fpz, Oz, Pz, Cz, P4, P3, P8, P7
11	P8, Oz, Fpz, Pz, Cz, P3, P4, P7	Fpz, Oz, P7, P4, P8, P3, Cz, Pz	Fpz, P7, Cz, P3, Pz, P4, Oz, P8
12	Pz, Fpz, Oz, Cz, P7, P8, P4, P3	P8, Fpz, P3, Oz, Pz, Cz, P7, P4	Pz, Fpz, P4, P8, P3, Oz, Cz, P7
Sbj	g-Tec Helmet Electrodes Day 1	Electrodes Day 2	Electrodes Day 3
1	Cz, Pz, Fpz, P3, P4, Oz, P7, P8	Cz, P4, P3, P7, Fpz, Pz, P8, Oz	Pz, Cz, P4, Fpz, P7, P3, P8, Oz
2	Fpz, P3, Cz, P8, Pz, P7, P4, Oz	P8, Cz, Oz, Fpz, P7, Pz, P4, P3	Fpz, P3, Cz, P8, Pz, Oz, P7, P4
3	Cz, P7, Oz, Fpz, Pz, P4, P3, P8	P4, Cz, P3, Pz, Oz, P8, P7, Fpz	P4, Fpz, Pz, P7, P8, Cz, Oz, P3
4	P4, Pz, P8, Cz, Fpz, P3, Oz, P7	P4, P3, P8, Cz, Pz, Fpz, Oz, P7	P4, Cz, Fpz, P7, P8, Oz, P3, Pz
5	Fpz, P7, Pz, Oz, Cz, P3, P8, P4	Pz, P7, P4, Fpz, P3, Oz, Cz, P8	Pz, Cz, Oz, P8, Fpz, P3, P7, P4
6	P4, Cz, Pz, P3, P7, Fpz, Oz, P8	Cz, P3, P4, P8, P7, Fpz, Pz, Oz	Pz, P4, Cz, P7, P8, P3, Oz, Fpz
7	Pz, P3, P4, P8, Oz, P7, Cz, Fpz	Cz, P3, P8, P7, Oz, P4, Pz, Fpz	P3, Oz, P8, Cz, Pz, Fpz, P7, P4
8	Cz, Pz, P8, P7, Oz, P4, P3, Fpz	Cz, Fpz, Oz, Pz, P8, P4, P7, P3	Cz, P3, P7, P8, Oz, Fpz, P4, Pz
9	Pz, P3, Fpz, P4, Cz, P7, P8, Oz	Fpz, Oz, P8, P7, P4, P3, Cz, Pz	Cz, Fpz, P8, Pz, P7, P4, P3, Oz
10	Cz, Fpz, P8, P3, P4, Oz, Pz, P7	Oz, Fpz, P7, Pz, Cz, P4, P8, P3	Fpz, P3, P8, P4, Pz, Oz, P7, Cz
11	Pz, Cz, P4, Oz, P7, P3, P8, Fpz	P8, Cz, Pz, P3, Fpz, P7, P4, Oz	Cz, P7, P8, P3, Pz, P4, Fpz, Oz
12	Pz, P8, Fpz, P3, Oz, P4, Cz, P7	Cz, Fpz, P8, P4, P7, Pz, Oz, P3	Pz, Fpz, Oz, P8, P7, P3, P4, Cz

Table 4.7: Electrodes selected by BLDA-FS for intra-subject analysis and cross-validation 12 (11 runs used as training set and 1 run used as validation set).

As in the previous case, the use of the Levenshtein distance is proposed to evaluate the differences between the selected electrodes from one day to another. The result of this analysis can be seen in the Figure 4.20. These results show that in all cases there are variations of the selected electrodes from one day to another, indicating that there is a difference within the same subject on the different days of the experiment.

It might be said that when a subject is observed to maintain a similar value of Levenshtein distance, and this is relatively low, the electrode structures are preserved without showing very great changes from one day to the next. However, in cases where the calculated distance values vary greatly, it can be said that the electrodes are not preserved from one day to the next, and in this case, it would be more difficult to generalize an electrode structure for more than one experimental day.

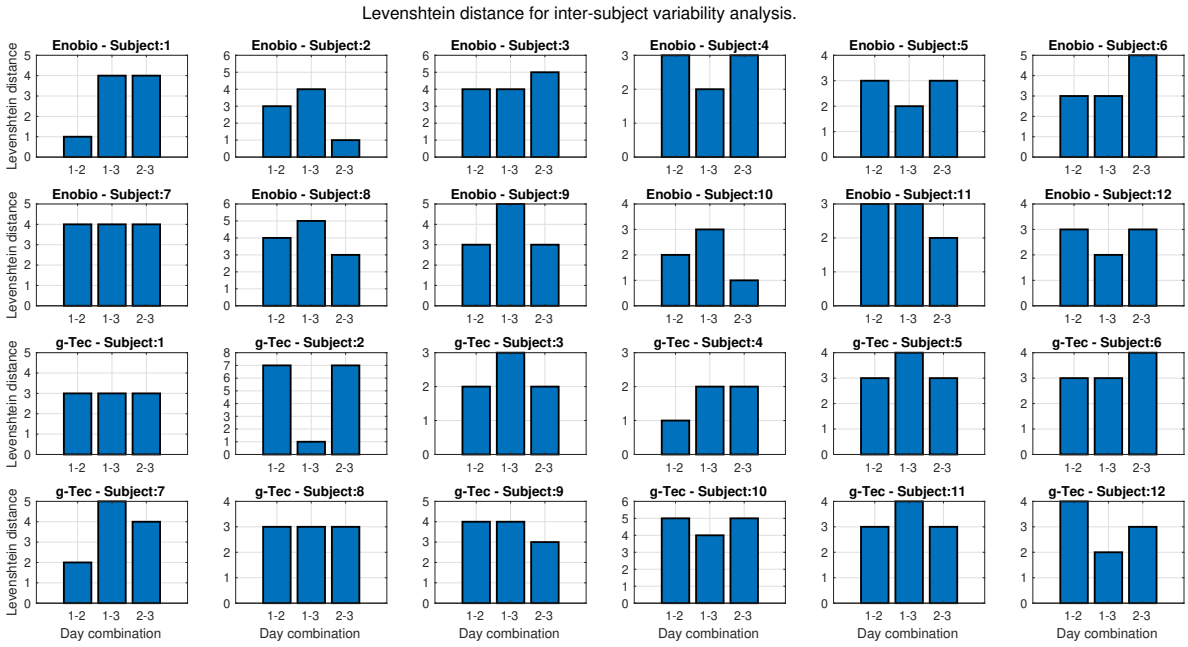


Figure 4.20: Levenshtein Distance results for intra-subject variability using cross-validation 12 (11 runs used as training set and 1 run used as validation set).

4.2.4 Can we generalize the model for different days?

This session will analyze whether it is possible to generalize the model over several days. Standard electrodes and a selection of electrodes using BLDA-FS will be used for this purpose.

4.2.4.1 BLDA-FS electrodes for model generalization

One of the questions this variability study attempts to answer whether it is possible to generalize a model for a subject. For this analysis, a selection of electrodes was proposed based on data from two experimental days and then classifying the third day using the selected electrodes to observe the accuracy results obtained.

Three cases were analyzed:

- Case A: Electrode selection from day 1 and 2 to be validated on day 3.
- Case B: Electrode selection from day 2 and 3 to be validated on day 1.
- Case C: Electrodes selected from day 1 and 3 to be validated on day 2.

For the selection of electrodes, in all cases there are 2 experimental days and each day has 2 sessions, i.e., there are 4 sessions. Each session consists of 6 runs, in this case there would be a total of 24 runs. It is proposed to cross-validate 24 as higher accuracy values have been obtained in previous cases. The system is trained with 23 runs and in each iteration, one is omitted for validation.

The accuracy results obtained in all cases are shown in Appendix B, Figure B.5 and Figure B.6. Also selected electrodes are shown in Table 4.8.

Enobio Helmet			
Sbj	Case A	Case B	Case C
1	Pz, Fpz, Oz, Cz, P8, P4, P3, P7	Cz, Fpz, P4, Pz, Oz, P7, P3, P8	Cz, Pz, Fpz, P3, P8, Oz, P7, P4
2	Pz, Fpz, Oz, Cz, P8, P4, P3, P7	Fpz, Oz, P7, P4, Cz, P3, Pz, P8	Fpz, Oz, P7, P4, Cz, P3, Pz, P8
3	Fpz, Cz, P7, P3, Pz, P4, Oz, P8	Cz, P7, Fpz, Pz, P3, Oz, P4, P8	Cz, P8, Fpz, P4, Pz, Oz, P3, P7
4	P3, P7, Fpz, Cz, P8, Pz, P4, Oz	Fpz, Oz, P4, P3, P7, Cz, P8, Pz	Fpz, P7, P3, Pz, P8, Cz, P4, Oz
5	Fpz, Pz, P7, P8, Cz, Oz, P3, P4	Fpz, P3, P7, P8, Pz, Cz, Oz, P4	Fpz, P3, Pz, Oz, P8, Cz, P7, P4
6	Pz, Fpz, Cz, Oz, P8, P4, P3, P7	Pz, Cz, Fpz, P8, P7, Oz, P4, P3	Cz, P8, Pz, Fpz, Oz, P4, P3, P7
7	Cz, Oz, P3, P8, P4, Pz, Fpz, P7	Fpz, Oz, P8, P7, P4, Cz, Pz, P3	Fpz, P7, Oz, Cz, Pz, P4, P3, P8
8	Pz, P4, Cz, Oz, P8, P3, Fpz, P7	Pz, Oz, P4, Cz, P8, P3, P7, Fpz	Cz, Oz, Pz, P4, P8, Fpz, P7, P3
9	Fpz, Oz, P8, P7, Pz, Cz, P4, P3	Fpz, P8, Oz, Pz, Cz, P4, P7, P3	Fpz, Oz, Cz, P3, P8, P7, P4, Pz
10	Fpz, P8, Pz, Oz, P7, P3, P4, Cz	Fpz, Oz, Cz, Pz, P7, P3, P4, P8	Fpz, P8, Cz, P7, Oz, P4, P3, Pz
11	Cz, Oz, Fpz, P4, Pz, P8, P7, P3	Fpz, P7, P4, Cz, P3, Pz, P8, Oz	Cz, P7, Fpz, P4, Oz, Pz, P3, P8
12	Pz, Fpz, Oz, P3, Cz, P8, P4, P7	Fpz, P8, Pz, P4, P3, P7, Oz, Cz	Pz, Fpz, P4, P8, Oz, P3, Cz, P7

g-Tec Helmet			
Sbj	Case A	Case B	Case C
1	Cz, Fpz, P3, Pz, P7, P8, P4, Oz	Pz, Cz, P7, P4, P8, P3, Fpz, Oz	Cz, Pz, P8, P4, P7, Fpz, P3, Oz
2	P8, Fpz, P7, Cz, P4, Oz, P3, Pz	Fpz, P7, Cz, Oz, P8, P3, P4, Pz	Fpz, P3, Cz, P7, P4, P8, Oz, Pz
3	Cz, P7, P4, P3, Oz, Pz, P8, Fpz	P4, Fpz, Cz, Pz, P7, P3, Oz, P8	Pz, P3, P4, P7, Fpz, Cz, Oz, P8
4	Pz, Cz, P8, Oz, P7, P4, P3, Fpz	Cz, P4, Fpz, Pz, P8, P7, P3, Oz	Cz, P7, P8, Fpz, P4, Oz, Pz, P3
5	Fpz, Cz, Oz, P7, Pz, P3, P8, P4	Pz, Fpz, P4, P3, P7, Cz, Oz, P8	Pz, Fpz, P3, P8, P7, Oz, P4, Cz
6	Cz, P4, P8, Fpz, Pz, P7, Oz, P3	Cz, P3, P4, P7, Pz, Oz, Fpz, P8	P4, Cz, Pz, P8, Fpz, Oz, P3, P7
7	Pz, P3, P8, Fpz, P4, Cz, Oz, P7	Cz, P3, P7, Pz, P8, Oz, P4, Fpz	Pz, Oz, Fpz, P3, P8, P7, P4, Cz
8	Cz, Fpz, P8, P7, Oz, Pz, P4, P3	Cz, P3, P8, P4, P7, Fpz, Pz, Oz	Pz, Cz, P3, Oz, P7, P8, Fpz, P4
9	Cz, Pz, P8, Fpz, P7, P3, P4, Oz	Fpz, Cz, P8, Pz, P4, Oz, P3, P7	Cz, Pz, P7, P8, P4, Oz, P3, Fpz
10	Fpz, P8, P4, P7, Oz, P3, Pz, Cz	Fpz, Pz, P8, Oz, P4, P7, P3, Cz	Fpz, P8, P4, Pz, P3, Cz, Oz, P7
11	Cz, Fpz, Oz, P7, P4, P8, Pz, P3	Cz, P8, Oz, P4, P7, Pz, Fpz, P3	Cz, Pz, P4, P7, Fpz, P3, P8, Oz
12	Pz, Fpz, P4, P8, Cz, P7, Oz, P3	Fpz, Pz, P4, Oz, P8, P3, P7, Cz	Pz, Fpz, P3, Oz, P8, P7, P4, Cz

Table 4.8: Electrodes selected by BLDA-FS for Case A, B and C.

After selecting the electrodes with the BLDA-FS algorithm, optimal electrodes are defined, i.e., electrodes with which the maximum accuracy can be obtained. The number of optimal electrodes for each subject and each case are shown in the following Table 4.9. From the results shown in this table, the number of electrodes considered optimal is higher than those shown in Tables 4.4 and 4.6.

Day	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	S11	S12	Mean	Variance
Enobio														
Case A	5	6	6	6	3	8	4	7	4	3	6	4	5.17	2.51
Case B	5	6	4	6	5	6	8	2	7	5	3	4	5.08	2.81
Case C	4	5	3	2	3	4	5	4	6	2	4	4	3.83	1.42
g-Tec														
Case A	6	5	4	3	4	5	3	3	5	6	8	7	4.91	2.63
Case B	5	7	4	3	3	5	4	5	6	4	3	4	4.42	1.54
Case C	6	4	7	3	5	5	5	7	7	3	7	4	5.25	2.38

Table 4.9: Optimal number of electrodes for cases A, B and C.

However, a summary of accuracy results is presented in Figure 4.21. After using these data for the validation process, the accuracy values shown in the Figure 4.22 were obtained:

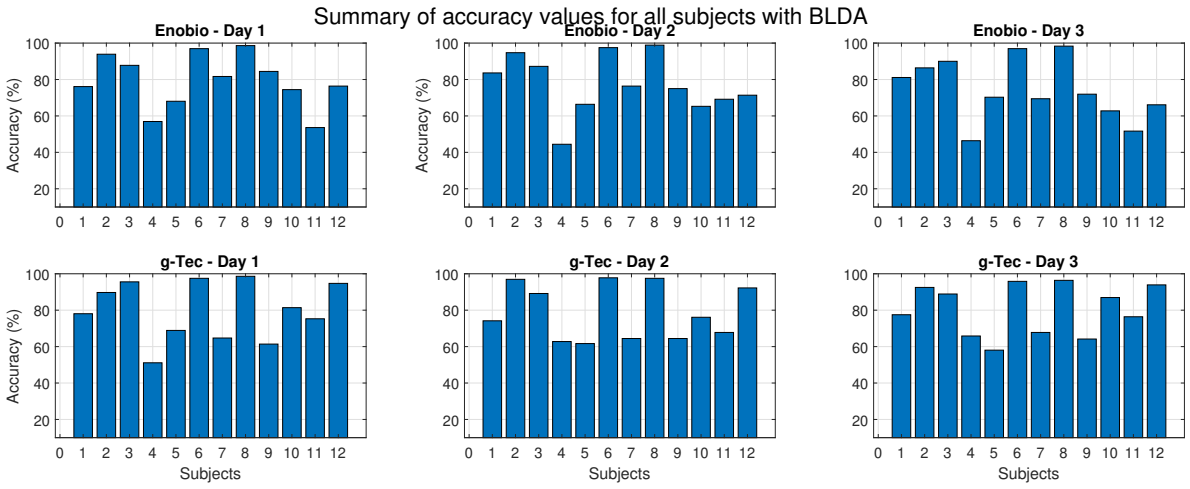


Figure 4.21: Accuracy values from the two days of Cases A, B and C for Enobio and g-Tec helmets.

Figure 4.23 shows the comparison between the accuracy results of the two days and the accuracy results of the last day. It can be clearly seen that the subjects who showed high accuracy in the previous analyses, e.g., subjects 2, 3, 6, and 8, continue to show high accuracy scores. On the other hand, the subjects who showed low accuracy values in the previous analyses show lower accuracy values after this analysis, e.g., subjects 4, 10, and 11 for the Enobio helmet and 5, 7, and 9 for the g-Tec helmet. From this figure, it can be seen how the error is higher when the model is to be generalized using parameters generated from unknown data.

Finally, the mean accuracy for all subjects is calculated and shown in Figure 4.24. The T-test was calculated over this result and showed a p-value of 0.0034 in the case of the Enobio helmet and a p-value of 0.0447 in the case of the g-Tec helmet. In both cases, the null hypothesis is rejected which shows that there is an important difference between the results. This means that there is a high variability between subjects and due to the existing difference it is difficult to generalize a model based

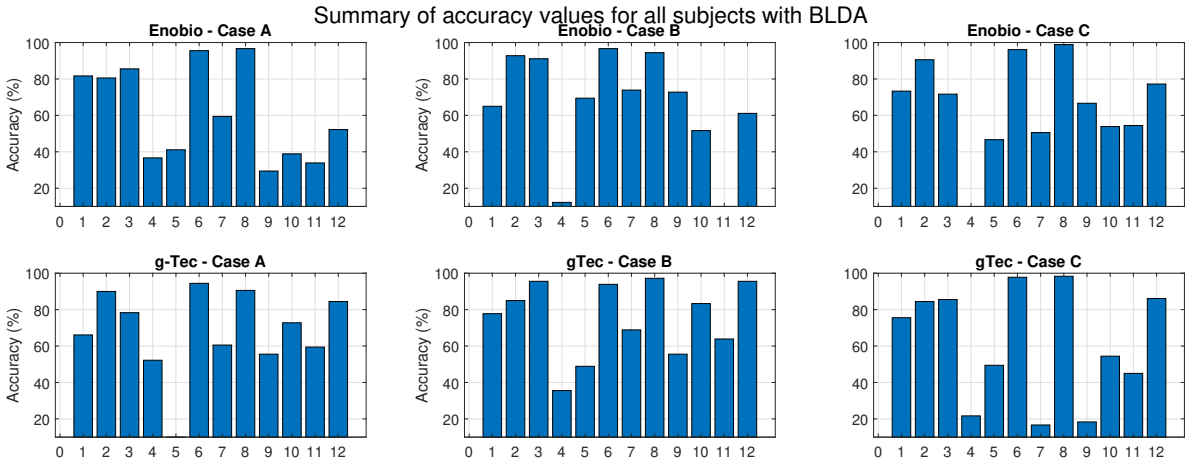


Figure 4.22: Accuracy results of Cases A, B and C for Enobio and g-Tec helmets using optimal BLDA-FS electrodes.

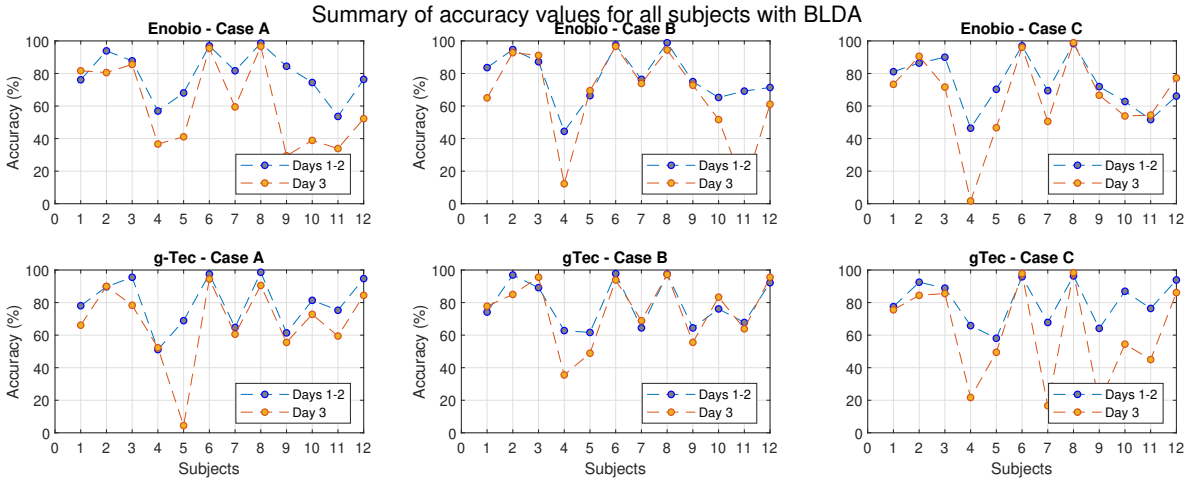


Figure 4.23: Accuracy results comparison of Cases A, B and C for Enobio and g-Tec helmets using optimal BLDA-FS electrodes.

on only part of the data.

The accuracy average shown in Figure 4.24 for 3 days of the omitted dataset (one day left) is 64.12% and 68.87% for the Enobio and g-Tec helmet, respectively. It can be observed that with two days of a subject and fewer electrodes, the performance is better than the case where all the information of all the subjects is considered, as shown in Figure 4.6.

4.2.4.2 Standard Electrodes for model generalization

For this final analysis, the cases described in the previous analysis are used, but this time the accuracy calculation is performed using standard electrodes. The Figure 4.25 shows the comparison between the accuracy results of the two days and the accuracy results of the last day. Figure 4.26 shows the average results for both helmets.

The results calculated with standard electrodes provide higher accuracy values than in the previ-

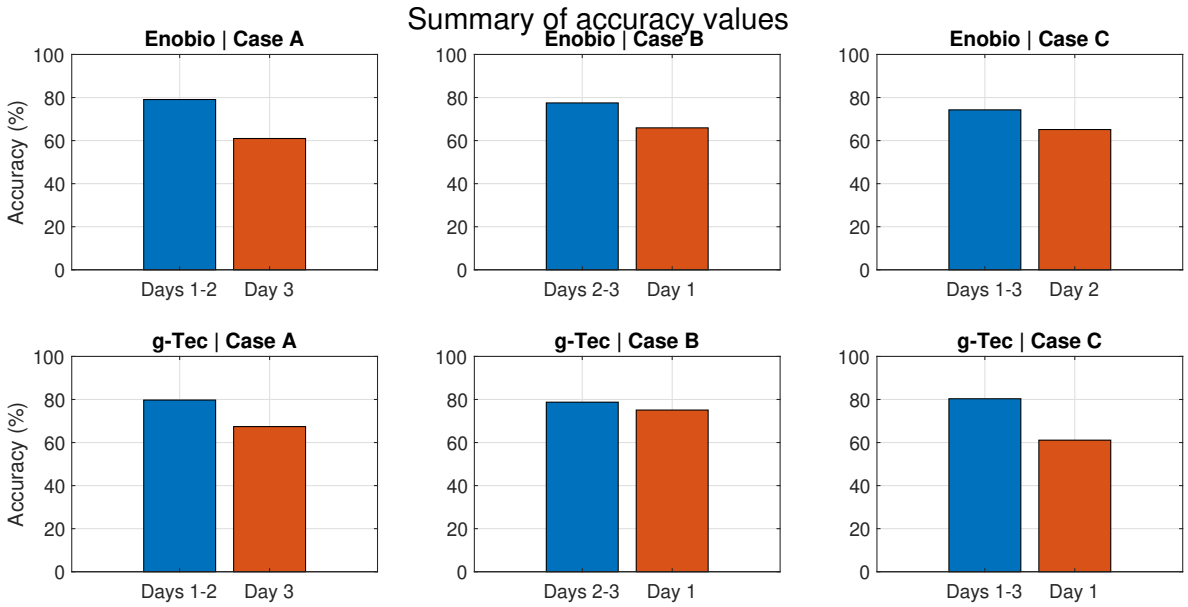


Figure 4.24: Comparison between the accuracy means of cases A,B and C with optimal BLDA-FS electrodes.

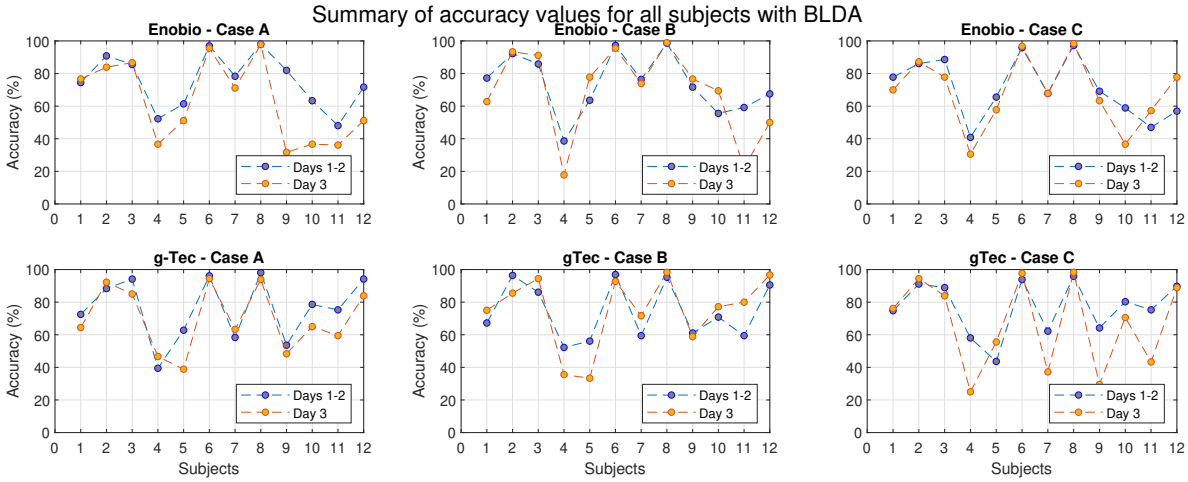


Figure 4.25: Accuracy results comparison of Cases A, B and C for Enobio and g-Tec helmets using standard electrodes.

ously analyzed case using the BLDA-FS electrodes. This suggests that the standard electrodes are better at generalizing on one experimental day than the electrodes selected by BLDA-FS for the other two days. The differences that exist between days suggest that the electrodes selected for the two days do not necessarily perform well on the third day. An exhaustive selection of optimal electrodes for a given data set was performed. This may result in overfitting the model for a given day, in this case two days, limiting the generalization and reducing the degrees of freedom of the model to fit unknown data.

A study with similar results can be seen in Figure 10 of (Changoluisa et al., 2020), where it is clearly observed that the standard electrodes provide better accuracy than the selected electrodes from an unknown dataset. The accuracy average is shown in Figure 4.26 for three days is 66.82% and 70.43% for Enobio and g-Tec helmet respectively.

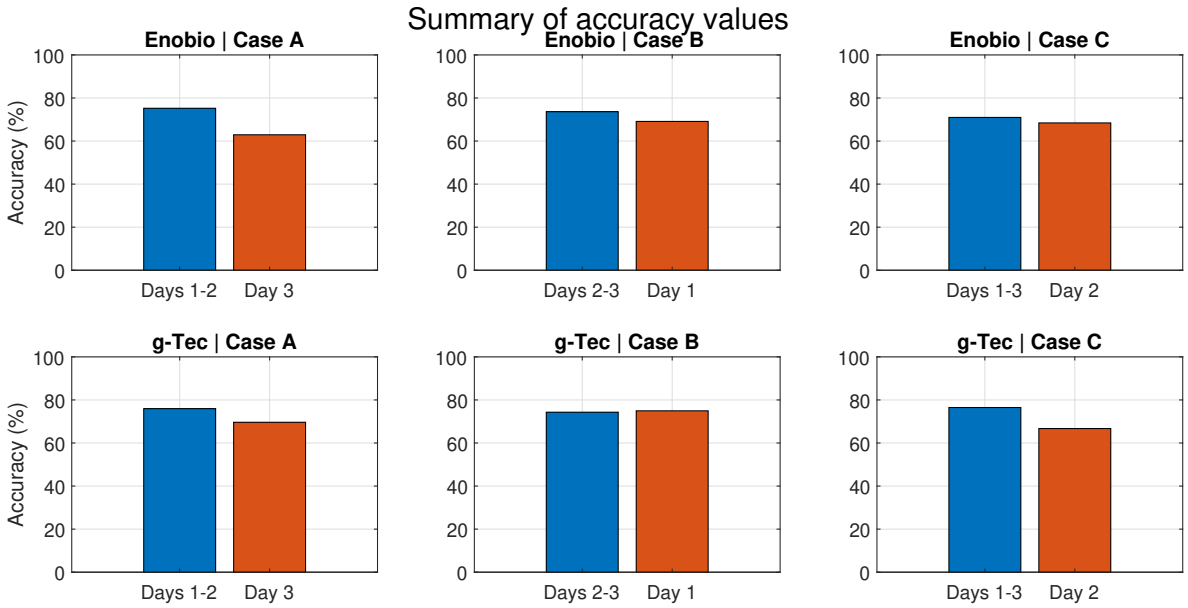


Figure 4.26: Comparison between the accuracy means of cases A,B and C with standard electrodes.

4.3 Summary of Results

A summary of results is presented in Table 4.10. This table summarizes the results of the preliminary investigations carried out. First, it should be noted that when a general analysis is performed, the accuracy values are the lowest of all the other cases. If the individual characteristics of the subjects are not accounted for, important information will be omitted due to the variability that the subjects exhibit. Despite the low results observed, it is important to highlight the difference between the two proposed cases of cross-validation. In the first case, cross-validation 12, in which one of the subjects is completely omitted, a lower accuracy is obtained than in the second case, cross-validation 72, in which almost all the information of all the subjects is taken into account (except one run), which allows the system to take into account the characteristics of all the data involved, increases the accuracy by about 5%.

In search of a deeper analysis, an inter-subjects analysis is proposed. In these results, it is observed that when using standard electrodes with an individual classification for each subject, the accuracy values are higher than in the general case, but they are lower than in the second consideration performed, where the optimal electrodes are selected for each subject using BLDA-FS. In both considerations, it can be observed that the cross-validations, which include the largest amount of information from all subjects, provide the highest accuracy results.

In addition, an intra-subject analysis was performed that not only attempts to evaluate the behavior of the subjects individually, but also to evaluate the behavior of the same subject in 3 experimental days and analyze how it varies. As in the previous cases, an analysis with standard electrodes is proposed first. In this case, it is observed that the results obtained are greater than in the case of inter-subject analysis, since the classification is focused on each data set belonging to each day. of the experiment. Then, an analysis is performed using selected electrodes with BLDA-FS, this analysis gives the best results among the other mentioned analyzes, since not only the classification is adapted to each experimental day of each subject, but also the electrodes selected for each subject are those that maximize the accuracy results.

Finally, to observe if it is possible to generalize the system between days, an analysis is performed considering 3 cases where two experimental days are used to select electrodes that are later used to validate the unknown day for the system. Along with this study, it is proposed to perform the same analysis but in a second approach using standard electrodes. From this analysis, it appears that the results obtained with the first approach (electrode selection with BLDA-FS) exceed those obtained with the general analysis (standard electrodes). This result indicates that it is possible to obtain good performance without having a large amount of data, as in the general analysis. With less data, it is possible to obtain similar accuracy results to the general analysis, as mentioned. On the other hand, if the results obtained with the standard electrodes are taken into account, it can be seen that the results are better from this approach.

When evaluating a dataset from parameters (electrodes) calculated based on the characteristics of another dataset unknown to the first one, the results will most likely not reach a high accuracy value. However, if these data are analyzed from standard electrodes selected, as indicated in Chapter 2, on the basis of many studies showing that these electrodes better capture certain characteristics of the signal; it is more likely that the results will be better than in the first case analyzed, as demonstrated in the study in (Changoluisa et al., 2020).

General Analysis		
Approach	Enobio Accuracy	g-Tec accuracy
Standard - Cross-validation 12	57.93%	62.21%
Standard - Cross-validation 72	62.3%	68.75%
Inter-subject Analysis		
Approach	Enobio Accuracy	g-Tec accuracy
Standard - Cross-validation 6	73.8%	77.22%
BLDA-FS - Cross-validation 6	75.88%	78.46%
Intra-subject Analysis		
Approach	Enobio Accuracy	g-Tec accuracy
Standard - Cross-validation 2	61.12%	64.61%
Standard - Cross-validation 12	66.82%	70.43%
BLDA-FS - Cross-validation 2	72.02%	75.99%
BLDA-FS - Cross-validation 12	76.10%	79.34%
Can we generalize the model?		
Approach	Enobio Accuracy	g-Tec accuracy
BLDA-FS - Cross-validation 24	64.12%	68.87%
Standard - Cross-validation 24	66.82%	70.42%

Table 4.10: Results summary.

5

Conclusions and Discussion

Due to the importance of the P300-ERP as a key element in the interpretation of brain activity during a cognitive process (van Dinteren et al., 2014), this project has attempted to elucidate how its properties vary as a function of different factors related to the subject and the experimental features used to obtain the signals. In order to analyze more clearly what was uncovered by the proposed project, an experimental process was conducted with the aim of eliciting the P300 signal through the oddball paradigm under the conditions explained in Section 1 of Chapter 3.

The results of the experiment were exhaustively analyzed to identify the individual characteristics of the 12 subjects during the experiment and to analyze the variability that can be derived from the data. Four of these experiments were conducted for this project, using two helmets: Enobio 8 and g-Tec g. USBAmp, both with active and passive dry electrodes, respectively.

As a first approximation, a preliminary examination of the data is proposed in Section 1 of Chapter 4, for which the measure r2 is used. From this analysis it appears that there is a clear differentiation of the P300 signal for most subjects, this is shown in Figures 4.1 and 4.2. In addition, a comparative analysis between subjects and helmets is proposed, the results of which are shown in Figure 4.4. From this analysis, it is found that there is a great similarity between the data of the same subjects in both cases, but they are not identical. It is also found that certain subjects have low similarity with respect to the other subjects, such as subjects 2, 8, 6, and 12.

From this analysis, as explained above, it was possible to identify a characteristic behavior or pattern of behavior in certain subjects that allows them to be identified among the others, suggesting that the individual characteristics observed in the subjects may be a ”fingerprint” that facilitates the recognition of the subject, noting that the P300 ERP is an important signature in cognitive processes (Linden, 2005).

The results of this analysis are consistent with the results of a second proposed exploratory analysis using the BLDA algorithm to classify the data from each subject with both helmets. The Figure 4.5 shows the obtained results indicating that subjects 2, 6, and 8 have the highest accuracy in both cases, in addition to subject 12 who also has high calculated accuracy values.

Through this analysis, it could be determined that the 3 subjects mentioned above provide the best

accuracy results. It was also found that there are relevant differences between the different subjects as well as between different helmets, supporting the theory of large variability between subjects, within a subject, and between helmets.

To analyze the observed differences in more detail, a variability analysis with two helmets of dry electrodes under certain aspects is proposed in Section 2.1 of Chapter 4. Therefore, in this project we studied the variability using the different approaches that we describe in more detail below.

General Analysis approach

First, a general analysis is performed with standard electrodes in both helmets, classifying all unified data without distinguishing subjects or experimental days. In this first result using cross-validation 12, where the information of one of the subjects is completely excluded, accuracy values of about 60% are obtained, as shown in Figure 4.6. In order to analyze how the obtained result changes depending on the information considered for the classification, a performance of a cross-validation 72 considering the information of all the subjects is proposed, improving the previously obtained results by about 5% as shown in Figure 4.7.

This shows that accuracy increases when the classification is adjusted to the characteristics of all subjects that are part of the dataset. Excluding all the information of a subject in the cross-validation reduces the performance of the classification because not all information is available, and the individual characteristics of that subject are not taken into account. However, this suggests that generalization for P300 detection with data from other subjects is a complicated issue.

From this analysis, the accuracy results for both helmets under similar characteristics of the experiment and analysis show no notable difference in P300 detection. This difference is about 5% in both cases. However, due to the differences between the helmets used, such as the size and spatial distribution of the electrodes, it is difficult to make a true comparison and it cannot be expected that the similarities or differences between the results obtained with both helmets are invariant.

Inter-subject Variability approach

Another aspect that should be evaluated is the difference that exists between subjects, that is, the inter-subject variability. To this end, a similar analysis was proposed in Section 2.2 of Chapter 4, which consists of determining the accuracy values for each subject independently using standard electrodes. Following this analysis, it was found that although the standard electrodes provided good results for some subjects such as subjects 1, 2, 3, 6, 8, and 9 for the Enobio helmet and subjects 2, 3, 6, 8, 10, and 12 for the g-Tec helmet, as shown in Figure 4.8, but did not provide good accuracy results in all cases, suggesting that to maximize accuracy, it is necessary to perform an individual study for each subject to determine the electrodes that best detect the analysis signal.

To this end, as in the previous case, a selection of electrodes was proposed by BLDA-FS. After this analysis, it is found that the accuracy values calculated with these electrodes are higher, as shown in Figures 4.10 and 4.11. Moreover, an optimal number of electrodes has been determined, which is smaller than 8 in almost all cases, to achieve maximum accuracy and optimize the system, as shown in Table 4.2.

This analysis has also made it possible to analyze the electrodes that are repeated the most in each case. In the Enobio helmet, the main electrodes are Cz and Fpz, while in the g-Tec helmet it is the Pz electrode. Electrodes from the frontal zone of the brain have shown great importance in (Changoluisa et al., 2020), electrode Fz represent this involvement.In both cases, a large participation of the parietal

lobe electrodes is observed, which is consistent with the literature highlighting the importance of these electrodes in maximizing recognition accuracy.

In the present analysis, it is important to emphasize that considerations must be made to determine the importance of an electrode. Each electrode must be normalized by assigning it an individual weight appropriate to the conditions of the electrodes. Factors considered important in normalizing electrodes include: rank order, which is the order in which electrodes appear; repetition, which is the number of times an electrode is repeated compared to other subjects; and the relative accuracy value each electrode provides. These three factors could define the weight that could be given to each electrode to evaluate its importance more objectively, this should be considered for future studies.

The results of the general analysis allowed to evaluate the general performance at the level of helmets from a different perspective. In this case the results obtained increase by about 10% with respect to the previously analyzed cases, but the behavior pattern is maintained: the difference between both helmets is not clear, and the g-Tec helmet shows slightly higher than the Enobio helmet.

Intra-subject variability approach

Finally, in Chapter 4, Section 2.2, an analysis of the differences that exist in the same subject on the different experimental days is proposed, i.e., intra-subject variability analysis. The experiment consisted of 3 experimental days with two sessions each, which will be analyzed individually. In a first approximation, standard electrodes are used for classification.

From this analysis, it can be highlighted that, as observed in the previous case, this choice of electrodes is useful for certain subjects, while for others it reduces accuracy, as it does not adapt to their individual characteristics, as shown in Figures 4.13 and 4.15, suggesting that adaptation of the methods used to evaluate the P300 signal is necessary for each individual subject as explained in (Changoluisa et al., 2018, Changoluisa et al., 2020). It was quite possible to observe, after an independent evaluation of each experimental day, that there were differences between the behaviors of the same subject on different days, even though the experimental conditions were very similar.

Comparing these results with the general analysis, we find that on certain days and for certain subjects the accuracy values are improved by these results, but this cannot be generalized for every due to the individual differences of each day and subject.

To adapt the evaluation method for the P300 signal, as in the previous cases, a selection with BLDA-FS is proposed, which allows the selection of the electrodes that best adapt to each experimental day. To this end, two considerations were made. The first consists of a cross-validation 2 in which one session is used for training and the other for testing. Second, cross-validation 12 is proposed in which 11 runs are considered for testing and only one for validation. In both cases, higher accuracy values are observed compared to the first case analyzed with standard electrodes as shown in Figures 4.17 and 4.19. However, in the second case, where practically all runs are considered for training, a significant improvement in accuracy results is observed. This because in this case most of the available data is learned and the selected electrodes consider most of the features of both sessions per day.

After this analysis, the difference between the selected electrodes for each day was analyzed using the Leveshtein distance, which, although it is a rather strict measure, allows us to quantify these differences. The difference between the different days was found to be very large and in many cases reached the maximum number of possible changes between electrodes, as shown in Figures 4.18 and 4.20. This result proves once again the large variability that can be observed even in the same subject performing the same task at different times.

This analysis has also made it possible to evaluate the behavior of the helmets on the different days, obtaining similar results to those obtained previously as shown in Figures 4.14 and 4.16. For day 1 and 3, the accuracy values of the g-Tec helmet are still slightly higher than in the case of the Enobio helmet. On day 2, the Enobio helmet is slightly higher. However, the differences are not remarkable in any case.

After analyzing the results obtained, it can be stated that variability exists between subjects and within the same subject and has been demonstrated and characterized with the analyzes performed, but it is important to consider other experimental characteristics that may affect the results obtained, such as the case of the electrodes used for the experiments. While it is true that dry electrodes offer some of the advantages described in Chapter 2, they also show certain limitations compared to wet electrodes. Comparing the obtained results with similar studies measured with wet electrodes, such as (Hoffmann et al., 2008, Changoluisa et al., 2020), it can be observed that the maximum accuracy values are lower in comparison, which could be due to the electrodes used. Other electrode selection methods can be used, such as the one proposed in Changoluisa et al., 2018, which has been shown to improve accuracy while accounting for inter- and intra-subject variability to analyze how results change.

Approach of the generalization of the model between days

As a final analysis, it was proposed in Chapter 4, Section 2.4 to perform an analysis of how the model can be generalized for different days. For this analysis, as a first approximation, a selection of electrodes using BLDA-FS over two experimental days is proposed to be evaluated on the third day, the results of this analysis are shown in Figure 4.23. The results show that when evaluating a third day, from selected electrodes of two unknown days, the accuracy obtained is 64.12% and 68.87% is obtained for Enobio and g-Tec helmets, respectively. In a second approximation, the same analysis is performed, in this case using standard electrodes. The results of this analysis of Figure 4.25, show that a higher accuracy can be obtained when using standard electrodes (66.82% and 70.42% for Enobio and g-Tec helmets, respectively) than when using the electrodes selected from an unknown data set. A possible explanation for the observed results is explained in Section 2.4 of Chapter 4.

Bibliography

Alexander, J. E., Bauer, L. O., Kuperman, S., Morzorati, S., O'Connor, S. J., Rohrbaugh, J., Porjesz, B., Begleiter, H., & Polich, J. (1996). Hemispheric differences for p300 amplitude from an auditory oddball task. *International Journal of Psychophysiology*, *21*(2), 189–196. https://doi.org/10.1016/0167-8760(95)00047-X

Alexander, J. E., Porjesz, B., Bauer, L. O., Kuperman, S., Morzorati, S., O'connor, S. J., Rohrbaugh, J., Begleiter, H., & Polich, J. (1995). P300 hemispheric amplitude asymmetries from a visual oddball task. *Psychophysiology*, *32*(5), 467–475. https://doi.org/10.1111/j.1469-8986.1995.tb02098.x

Allison, B. Z., Wolpaw, E. W., & Wolpaw, J. R. (2007). Brain–computer interface systems: Progress and prospects. *Expert Review of Medical Devices*, *4*(4), 463–474. https://doi.org/10.1586/17434440.4.4.463

Alwin, D. F. (1994). Aging, personality, and social change: The stability of individual differences over the adult life span. *Life-span development and behavior, vol. 12* (pp. 135–185). Lawrence Erlbaum Associates, Inc.

Aranzana, M. J., Garcia[U+2010]Mas, J., Carbó, J., & Arús, P. (2002). Development and variability analysis of microsatellite markers in peach. *Plant Breeding*, *121*(1), 87–92. https://doi.org/https://doi.org/10.1046/j.1439-0523.2002.00656.x

Bassett, D. S., Wymbs, N. F., Porter, M. A., Mucha, P. J., Carlson, J. M., & Grafton, S. T. (2011). Dynamic reconfiguration of human brain networks during learning. *Proceedings of the National Academy of Sciences*, *108*(18), 7641–7646. https://doi.org/10.1073/pnas.1018985108

Baykara, E., Ruf, C. A., Fioravanti, C., Käthner, I., Simon, N., Kleih, S. C., Kübler, A., & Halder, S. (2016). Effects of training and motivation on auditory p300 brain–computer interface performance. *Clinical Neurophysiology*, *127*(1), 379–387. https://doi.org/10.1016/j.clinph.2015.04.054

Blankertz, B., Lemm, S., Treder, M., Haufe, S., & Müller, K.-R. (2011). Single-trial analysis and classification of ERP components — a tutorial. *NeuroImage*, *56*(2), 814–825. https://doi.org/10.1016/j.neuroimage.2010.06.048

BrainSecrets. (n.d.). *Enobio NE* [Wireless quantitative EEG system]. Retrieved February 19, 2021, from http://www.brainsecrets.co.kr/default/menu03/menu03_prod08.php?sub=32

Bravi, A., Longtin, A., & Seely, A. J. (2011). Review and classification of variability analysis techniques with clinical applications. *BioMedical Engineering OnLine*, *10*(1), 90. https://doi.org/10.1186/1475-925X-10-90

Brouwer, A.-M., & Van Erp, J. B. F. (2010). A tactile p300 brain-computer interface. *Frontiers in Neuroscience*, *4*. https://doi.org/10.3389/fnins.2010.00019

Brunner, P., Joshi, S., Briskin, S., Wolpaw, J., Bischof, H., & Schalk, G. (2010). Does the "p300" speller depend on eye gaze? *Journal of neural engineering*, *7*(5), 056013. https://doi.org/10.1088/1741-2560/7/5/056013

Cao, B., Han, X., Tang, J., Zhou, Z., & Liu, Y. (2018). An optimizational tactile p300 brain-computer interface paradigm. *2018 37th Chinese Control Conference (CCC)*, 8215–8220. https://doi.org/10.23919/ChiCC.2018.8484236

Cecotti, H., Rivet, B., Congedo, M., Jutten, C., Bertrand, O., Maby, E., & Mattout, J. (2011). A robust sensor-selection method for p300 brain–computer interfaces. *Journal of Neural Engineering*, *8*(1), 016001. https://doi.org/10.1088/1741-2560/8/1/016001

Chan, H.-L., Kuo, P.-C., Cheng, C.-Y., & Chen, Y.-S. (2018). Challenges and future perspectives on electroencephalogram-based biometrics in person recognition [Publisher: Frontiers]. *Frontiers in Neuroinformatics*, *12*. https://doi.org/10.3389/fninf.2018.00066

Chandra, B. (2016, January 1). Chapter 3 - gene selection methods for microarray data. In D. Al-Jumeily, A. Hussain, C. Mallucci, & C. Oliver (Eds.), *Applied computing in medicine and health* (pp. 45–78). Morgan Kaufmann. https://doi.org/10.1016/B978-0-12-803468-2.00003-5

Changoluisa, V. (n.d.). *Towards smarter brain computer interface (BCI) – study of electroencephalographic signal processing and classification techniques towards the use of BCI intelligent and adaptive* (Doctoral dissertation).

Changoluisa, V., Varona, P., & Rodríguez, F. B. (2018). An electrode selection approach in p300-based BCIs to address inter- and intra-subject variability. *2018 6th International Conference on Brain-Computer Interface (BCI)*, 1–4. https://doi.org/10.1109/IWW-BCI.2018.8311497

Changoluisa, V., Varona, P., & Rodríguez, F. B. (2020). A low-cost computational method for characterizing event-related potentials for BCI applications and beyond. *IEEE Access*, *8*, 111089–111101. https://doi.org/10.1109/ACCESS.2020.3000187

Christensen, H., Mackinnon, A. J., Korten, A. E., Jorm, A. F., Henderson, A. S., & Jacomb, P. (1999). Dispersion in cognitive ability as a function of age: A longitudinal study of an elderly community sample. *Aging, Neuropsychology, and Cognition*, *6*(3), 214–228. https://doi.org/10.1076/anec.6.3.214.779

Cincotti, F., Kauhanen, L., Aloise, F., Palomäki, T., Caporusso, N., Jylänki, P., Mattia, D., Babiloni, F., Vanacker, G., Nuttin, M., Marciani, M. G., & Millán, J. d. R. (2007, September 2). *Vibrotactile feedback for brain-computer interface operation* [Computational intelligence and neuroscience]. https://doi.org/https://doi.org/10.1155/2007/48937

Cobb, G. W., & Moore, D. S. (1997). Mathematics, statistics, and teaching. *The American Mathematical Monthly*, *104*(9), 801–823. https://doi.org/10.1080/00029890.1997.11990723

Cole, M. W., Bassett, D. S., Power, J. D., Braver, T. S., & Petersen, S. E. (2014). Intrinsic and task-evoked network architectures of the human brain. *Neuron*, *83*(1), 238–251. https://doi.org/10.1016/j.neuron.2014.05.014

Comerchero, M. D., & Polich, J. (1999). P3a and p3b from typical auditory and visual stimuli. *Clinical Neurophysiology*, *110*(1), 24–30. https://doi.org/10.1016/S0168-5597(98)00033-1

Damon, W., Lerner, R. M., Kuhn, D., & Siegler, R. S. (2006, May 11). *Handbook of child psychology, cognition, perception, and language*. John Wiley & Sons.

Demuru, M., & Fraschini, M. (2020). EEG fingerprinting: Subject-specific signature based on the aperiodic component of power spectrum. *Computers in Biology and Medicine*, *120*, 103748. https://doi.org/10.1016/j.compbiomed.2020.103748

Dixon, D. o. P. R., Dixon, R. A., Bäckman, L., Park, D. C., Backman, P. o. P. a. U. o. G. D. o. P. S. D. o. N. bibinitperiod G. M. L., & Nilsson, L.-G. (2004). *New frontiers in cognitive aging*. Oxford University Press.

Donchin, E., Spencer, K., & Wijesinghe, R. (2000). The mental prosthesis: Assessing the speed of a p300-based brain-computer interface. *IEEE Transactions on Rehabilitation Engineering*, *8*(2), 174–179. https://doi.org/10.1109/86.847808

Donchin, E. (1981). Surprise!... surprise? *Psychophysiology*, *18*(5), 493–513. https://doi.org/10.1111/j.1469-8986.1981.tb01815.x

Donchin, E., & Coles, M. G. H. (1988). Is the p300 component a manifestation of context updating? *Behavioral and Brain Sciences*, *11*(3), 357–374. https://doi.org/10.1017/S0140525X00058027

Duncan-Johnson, C. C., & Donchin, E. (1977). On quantifying surprise: The variation of event-related potentials with subjective probability. *Psychophysiology*, *14*(5), 456–467. https://doi.org/10.1111/j.1469-8986.1977.tb01312.x

Emmerson, R. Y., Dustman, R. E., Shearer, D. E., & Turner, C. W. (1989). P3 latency and symbol digit performance correlations in aging. *Experimental Aging Research*, *15*(3), 151–159. https://doi.org/10.1080/03610738908259769

Farwell, L. A., & Donchin, E. (1988). Talking off the top of your head: Toward a mental prosthesis utilizing event-related brain potentials. *Electroencephalography and Clinical Neurophysiology*, *70*(6), 510–523.

Fazel-Rezai, R., Allison, B. Z., Guger, C., Sellers, E. W., Kleih, S. C., & Kübler, A. (2012). P300 brain computer interface: Current challenges and emerging trends. *Frontiers in Neuroengineering*, *5*, 14. https://doi.org/10.3389/fneng.2012.00014

Fazel-Rezai, R., Gavett, S., Ahmad, W., Rabbi, A., & Schneider, E. (2011). A comparison among several p300 brain-computer interface speller paradigms. *Clinical EEG and neuroscience*, *42*(4), 209–213. https://doi.org/10.1177/155005941104200404

Fitzgerald, P. G., & Picton, T. W. (1981). Temporal and sequential probability in evoked potential studies. *Canadian Journal of Psychology/Revue canadienne de psychologie*, *35*(2), 188–200. https://doi.org/10.1037/h0081154

Fraschini, M., Hillebrand, A., Demuru, M., Didaci, L., & Marcialis, G. L. (2015). An EEG-based biometric system using eigenvector centrality in resting state brain networks [Conference Name: IEEE Signal Processing Letters]. *IEEE Signal Processing Letters*, *22*(6), 666–670. https://doi.org/10.1109/LSP.2014.2367091

Furdea, A., Halder, S., Krusienski, D. J., Bross, D., Nijboer, F., Birbaumer, N., & Kübler, A. (2009). An auditory oddball (p300) spelling system for brain-computer interfaces. *Psychophysiology*, *46*(3), 617–625. https://doi.org/10.1111/j.1469-8986.2008.00783.x

González, R. C. (2008). Digital image processing [Edition: 3rd ed. ISBN: 9780135052679 Place: Upper Saddle River].

Gordon, B., & Carson, K. (1990). The basis for choice reaction time slowing in alzheimer's disease. *Brain and Cognition*, *13*(2), 148–166. https://doi.org/10.1016/0278-2626(90)90047-r

Graimann, B., Allison, B., & Pfurtscheller, G. (Eds.). (2010). *Brain-computer interfaces: Revolutionizing human-computer interaction*. Springer.

Grozea, C., Voinescu, C. D., & Fazli, S. (2011). Bristle-sensors—low-cost flexible passive dry EEG electrodes for neurofeedback and BCI applications. *Journal of Neural Engineering*, *8*(2), 025008. https://doi.org/10.1088/1741-2560/8/2/025008

Guger, C., Schlogl, A., Neuper, C., Walterspacher, D., Strein, T., & Pfurtscheller, G. (2001). Rapid prototyping of an EEG-based brain-computer interface (BCI). *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *9*(1), 49–58. https://doi.org/10.1109/7333.918276

Guger, C., Daban, S., Sellers, E., Holzner, C., Krausz, G., Carabalona, R., Gramatica, F., & Edlinger, G. (2009). How many people are able to control a p300-based brain-computer interface (BCI)? *Neuroscience Letters*, *462*(1), 94–98. https://doi.org/10.1016/j.neulet.2009.06.045

Guger, C., Spataro, R., Allison, B. Z., Heilinger, A., Ortner, R., Cho, W., & La Bella, V. (2017). Complete locked-in and locked-in patients: Command following assessment and communication with vibro-tactile p300 and motor imagery brain-computer interface tools. *Frontiers in Neuroscience*, *11*. https://doi.org/10.3389/fnins.2017.00251

Guger, C., Spataro, R., Pellas, F., Allison, B. Z., Heilinger, A., Ortner, R., Cho, W., Xu, R., La Bella, V., Edlinger, G., Annen, J., Mandalá, G., Chatelle, C., & Laureys, S. (2018). Assess-

ing command-following and communication with vibro-tactile p300 brain-computer interface tools in patients with unresponsive wakefulness syndrome. *Frontiers in Neuroscience*, *12*. https://doi.org/10.3389/fnins.2018.00423

Guo, C., Duan, L., Li, W., & Paller, K. A. (2006). Distinguishing source memory and item memory: Brain potentials at encoding and retrieval. *Brain Research*, *1118*(1), 142–154. https://doi.org/10.1016/j.brainres.2006.08.034

Guo, F., Hong, B., Gao, X., & Gao, S. (2008). A brain-computer interface using motion-onset visual evoked potential. *Journal of Neural Engineering*, *5*(4), 477–485. https://doi.org/10.1088/1741-2560/5/4/011

Halder, S., Rea, M., Andreoni, R., Nijboer, F., Hammer, E. M., Kleih, S. C., Birbaumer, N., & Kübler, A. (2010). An auditory oddball brain–computer interface for binary choices. *Clinical Neurophysiology*, *121*(4), 516–523. https://doi.org/10.1016/j.clinph.2009.11.087

Hale, S., Myerson, J., Smith, G. A., & Poon, L. W. (1988). Age, variability, and speed: Between-subjects diversity. *Psychology and Aging*, *3*(4), 407–410. https://doi.org/10.1037//0882-7974.3.4.407

Herrera, L., Rubio, G., Pomares, H., Paechter, B., Guillén, A., & Rojas, I. (2009). Strengthening the forward variable selection stopping criterion. *5769*, 215–224. https://doi.org/10.1007/978-3-642-04277-5_22

Hertzog, C., Dixon, R. A., & Hultsch, D. F. (1992). Intraindividual change in text recall of the elderly. *Brain and Language*, *42*(3), 248–269. https://doi.org/10.1016/0093-934X(92)90100-S

HFTCo., H. (n.d.). *G.SAHARAsys* [Smart technologist]. Retrieved February 19, 2021, from http://hooshmandfanavar.com/en/g-tec-medical-engineering-solutions/g-saharasys/

Hill, N. J., Lal, T. N., Bierig, K., Birbaumer, N., & Schölkopf, B. (n.d.). An auditory paradigm for brain-computer interfaces, 8.

Hillyard, S. A., Hink, R. F., Schwent, V. L., & Picton, T. W. (1973). Electrical signs of selective attention in the human brain. *Science (New York, N.Y.)*, *182*(4108), 177–180. https://doi.org/10.1126/science.182.4108.177

Ho, J. W., Stefani, M., dos Remedios, C. G., & Charleston, M. A. (2008). Differential variability analysis of gene expression and its application to human diseases. *Bioinformatics*, *24*(13), i390–i398. https://doi.org/10.1093/bioinformatics/btn142

Hoffmann, U., Vesin, J.-M., Ebrahimi, T., & Diserens, K. (2008). An efficient p300-based brain–computer interface for disabled subjects. *Journal of Neuroscience Methods*, *167*(1), 115–125. https://doi.org/10.1016/j.jneumeth.2007.03.005

Hong, B., Guo, F., Liu, T., Gao, X., & Gao, S. (2009). N200-speller using motion-onset visual response. *Clinical Neurophysiology*, *120*(9), 1658–1666. https://doi.org/10.1016/j.clinph.2009.06.026

Honjoh, Y., & Okita, T. (1981). [the effects of selective attention to tone pips and speech messages upon event-related potential (author's transl)]. *Shinrigaku Kenkyu: The Japanese Journal of Psychology*, *52*(1), 23–29.

Houlihan, M., Stelmack, R., & Campbell, K. (1998). Intelligence and the effects of perceptual processing demands, task difficulty and processing speed on p300, reaction time and movement time. *Intelligence*, *26*(1), 9–25. https://doi.org/10.1016/S0160-2896(99)80049-X

Hultsch, D. F., MacDonald, S. W., Hunter, M. A., Levy-Bencheton, J., & Strauss, E. (2000). Intraindividual variability in cognitive performance in older adults: Comparison of adults with mild dementia, adults with arthritis, and healthy adults. *Neuropsychology*, *14*(4), 588–598. https://doi.org/10.1037//0894-4105.14.4.588

Intriligator, J., & Polich, J. (1995). On the relationship between EEG and ERP variability. *International Journal of Psychophysiology: Official Journal of the International Organization of Psychophysiology*, *20*(1), 59–74. https://doi.org/10.1016/0167-8760(95)00028-q

Isreal, J. B., Chesney, G. L., Wickens, C. D., & Donchin, E. (1980). P300 and tracking difficulty: Evidence for multiple resources in dual-task performance. *Psychophysiology*, *17*(3), 259–273. https://doi.org/10.1111/j.1469-8986.1980.tb00146.x

Ito, J., Shibasaki, H., & Kimura, J. (1992). Somatosensory event-related potentials following different stimulus conditions. *The International Journal of Neuroscience*, *65*(1), 239–246. https://doi.org/10.3109/00207459209003297

Ivezić, Ž., Connolly, A. J., VanderPlas, J. T., & Gray, A. (2014, January 12). *Statistics, data mining, and machine learning in astronomy: A practical python guide for the analysis of survey data*. Princeton University Press.

Johnson, R., & Donchin, E. (1978). On how p300 amplitude varies with the utility of the eliciting stimuli. *Electroencephalography and Clinical Neurophysiology*, *44*(4), 424–437. https://doi.org/10.1016/0013-4694(78)90027-5

Kaper, M., Meinicke, P., Grossekathoefer, U., Lingner, T., & Ritter, H. (2004). BCI competition 2003-data set IIb: Support vector machines for the p300 speller paradigm. *IEEE Transactions on Biomedical Engineering*, *51*(6), 1073–1076. https://doi.org/10.1109/TBME.2004.826698

Käthner, I., Ruf, C. A., Pasqualotto, E., Braun, C., Birbaumer, N., & Halder, S. (2013). A portable auditory p300 brain–computer interface with directional cues. *Clinical Neurophysiology*, *124*(2), 327–338. https://doi.org/10.1016/j.clinph.2012.08.006

Kaufmann, T., Holz, E. M., & Kübler, A. (2013). Comparison of tactile, auditory, and visual modality for brain-computer interface use: A case study with a patient in the locked-in state. *Frontiers in Neuroscience*, *7*. https://doi.org/10.3389/fnins.2013.00129

Klobassa, D. S., Vaughan, T. M., Brunner, P., Schwartz, N. E., Wolpaw, J. R., Neuper, C., & Sellers, E. W. (2009). Toward a high-throughput auditory p300-based brain–computer interface. *Clinical Neurophysiology*, *120*(7), 1252–1261. https://doi.org/10.1016/j.clinph.2009.04.019

Krusienski, D. J., Sellers, E. W., Cabestaing, F., Bayoudh, S., McFarland, D. J., Vaughan, T. M., & Wolpaw, J. R. (2006). A comparison of classification techniques for the p300 speller. *Journal of Neural Engineering*, *3*(4), 299. https://doi.org/10.1088/1741-2560/3/4/007

Kübler, A., Nijboer, F., Mellinger, J., Vaughan, T. M., Pawelzik, H., Schalk, G., McFarland, D. J., Birbaumer, N., & Wolpaw, J. R. (2005). Patients with ALS can use sensorimotor rhythms to operate a brain-computer interface. *Neurology*, *64*(10), 1775–1777. https://doi.org/10.1212/01.WNL.0000158616.43002.6D

Kübler, A., Kotchoubey, B., Kaiser, J., Wolpaw, J. R., & Birbaumer, N. (2001). Brain–computer communication: Unlocking the locked in. *Psychological Bulletin*, *127*(3), 358–375. https://doi.org/10.1037/0033-2909.127.3.358

Levenshtein, V. I. (1966). Binary codes capable of correcting deletions, insertions and reversals. *Soviet Physics Doklady*, *10*, 707. Retrieved March 10, 2021, from http://adsabs.harvard.edu/abs/1966SPhD...10..707L

Li, F., Tao, Q., Peng, W., Zhang, T., Si, Y., Zhang, Y., Yi, C., Biswal, B., Yao, D., & Xu, P. (2020). Inter-subject p300 variability relates to the efficiency of brain networks reconfigured from resting- to task-state: Evidence from a simultaneous event-related EEG-fMRI study. *NeuroImage*, *205*, 116285. https://doi.org/10.1016/j.neuroimage.2019.116285

Li, S., Aggen, S. H., Nesselroade, J. R., & Baltes, P. B. (2001). Short-term fluctuations in elderly people's sensorimotor functioning predict text and spatial memory performance: The macarthur successful aging studies. *Gerontology*, *47*(2), 100–116. https://doi.org/10.1159/000052782

Linden, D. E. J. (2005). The p300: Where in the brain is it produced and what does it tell us? *The Neuroscientist: A Review Journal Bringing Neurobiology, Neurology and Psychiatry*, *11*(6), 563–576. https://doi.org/10.1177/1073858405280524

Lu, J., Speier, W., Hu, X., & Pouratian, N. (2013). The effects of stimulus timing features on p300 speller performance. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, *124*(2), 306–314. https://doi.org/10.1016/j.clinph.2012.08.002

Lu, S., Guan, C., & Zhang, H. (2008). Unsupervised brain computer interface based on inter-subject information. *2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, 638–641. https://doi.org/10.1109/IEMBS.2008.4649233

Makella, J., & Jousmaki, V. (2001). Magnetoencephalography. *Clinical Neurophysiology*, *112*, S26–S27. https://doi.org/10.1016/S1388-2457(01)80017-1

Matsuo, T., Iinuma, K., & Esashi, M. (1973). A barium-titanate-ceramics capacitive-type EEG electrode. *IEEE Transactions on Biomedical Engineering*, *BME-20*(4), 299–300. https://doi.org/10.1109/TBME.1973.324197

Meir-Hasson, Y., Kinreich, S., Podlipsky, I., Hendler, T., & Intrator, N. (2014). An EEG finger-print of fMRI deep regional activation. *NeuroImage*, *102 Pt 1*, 128–141. https://doi.org/10.1016/j.neuroimage.2013.11.004

Nam, C., Nijholt, A., & Lotte, F. (2014). *Brain-computer interfaces handbook: Technological and theoretical advances* (Second).

Nesselroade, J. R., & Featherman, D. L. (1997). Establishing a reference frame against which to chart age-related changes. *Studying aging and social change: Conceptual and methodological issues* (pp. 191–205). Sage Publications, Inc.

Nijboer, F., Furdea, A., Gunst, I., Mellinger, J., McFarland, D. J., Birbaumer, N., & Kübler, A. (2008). An auditory brain–computer interface (BCI). *Journal of Neuroscience Methods*, *167*(1), 43–50. https://doi.org/10.1016/j.jneumeth.2007.02.009

O'Donnell, B. F., Friedman, S., Squires, N. K., Maloon, A., & et al. (1990). Active and passive p3 latency in dementia: Relationship to psychometric, electroencephalographic, and computed tomographic measures. *Neuropsychiatry, Neuropsychology, & Behavioral Neurology*, *3*(3), 164–179.

O'Donnell, B. F., Squires, N. K., Martz, M. J., Chen, J.-R., & Phay, A. J. (1987). Evoked potential changes and neuropsychological performance in parkinson's disease. *Biological Psychology*, *24*(1), 23–37. https://doi.org/10.1016/0301-0511(87)90097-4

Okita, T. (1979). Event-related potentials and selective attention to auditory stimuli varying in pitch and localization. *Biological Psychology*, *9*(4), 271–284. https://doi.org/10.1016/0301-0511(79)90027-9

Ou, B., Wu, C., Zhao, G., & Wu, J. (2012). P300 amplitude reflects individual differences of navigation performance in a driving task. *International Journal of Industrial Ergonomics*, *42*(1), 8–16. https://doi.org/10.1016/j.ergon.2011.11.006

Ouyang, G., Hildebrandt, A., Sommer, W., & Zhou, C. (2017a). Exploiting the intra-subject latency variability from single-trial event-related potentials in the p3 time range: A review and comparative evaluation of methods. *Neuroscience & Biobehavioral Reviews*, *75*, 1–21. https://doi.org/10.1016/j.neubiorev.2017.01.023

Ouyang, G., Hildebrandt, A., Sommer, W., & Zhou, C. (2017b). Exploiting the intra-subject latency variability from single-trial event-related potentials in the p3 time range: A review and comparative evaluation of methods. *Neuroscience & Biobehavioral Reviews*, *75*, 1–21. https://doi.org/10.1016/j.neubiorev.2017.01.023

Palaniappan, R., Paramesran, R., Nishida, S., & Saiwaki, N. (2002). A new brain-computer interface design using fuzzy ARTMAP. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *10*(3), 140–148. https://doi.org/10.1109/TNSRE.2002.802854

Peng, H.-L., Liu, J.-Q., Tian, H.-C., Xu, B., Dong, Y.-Z., Yang, B., Chen, X., & Yang, C.-S. (2015). Flexible dry electrode based on carbon nanotube/polymer hybrid micropillars for biopotential recording. *Sensors and Actuators A: Physical*, *235*, 48–56. https://doi.org/10.1016/j.sna.2015.09.024

Perelmouter, J., & Birbaumer, N. (2000). A binary spelling interface with random errors. *IEEE Transactions on Rehabilitation Engineering*, *8*(2), 227–232. https://doi.org/10.1109/86.847824

Piccione, F., Giorgi, F., Tonin, P., Priftis, K., Giove, S., Silvoni, S., Palmas, G., & Beverina, F. (2006). P300-based brain computer interface: Reliability and performance in healthy and paralysed participants. *Clinical Neurophysiology*, *117*(3), 531–537. https://doi.org/10.1016/j.clinph.2005.07.024

Picton, T. (1992). The p300 wave of the human event-related potential. *Journal of Clinical Neurophysiology*, *9*(4), 456–479. Retrieved February 24, 2020, from insights.ovid.com

Polich, J. (1996). Meta-analysis of p300 normative aging studies. *Psychophysiology*, *33*(4), 334–353. https://doi.org/10.1111/j.1469-8986.1996.tb01058.x

Polich, J. (1997). On the relationship between EEG and p300: Individual differences, aging, and ultradian rhythms. *International Journal of Psychophysiology*, *26*(1), 299–317. https://doi.org/10.1016/S0167-8760(97)00772-1

Polich, J. (2007). Updating p300: An integrative theory of p3a and p3b. *Clinical neurophysiology : official journal of the International Federation of Clinical Neurophysiology*, *118*(10), 2128–2148. https://doi.org/10.1016/j.clinph.2007.04.019

Polich, J., Ladish, C., & Burns, T. (1990). Normal variation of p300 in children: Age, memory span, and head size. *International Journal of Psychophysiology*, *9*(3), 237–248. https://doi.org/10.1016/0167-8760(90)90056-J

Polikoff, J. B., Bunnell, H. T., & Jr, W. J. B. (n.d.). *Toward a p300-based computer interface*.

Popescu, F., Fazli, S., Badower, Y., Blankertz, B., & Müller, K.-R. (2007). Single trial classification of motor imagination using 6 dry EEG electrodes. *PLOS ONE*, *2*(7), e637. https://doi.org/10.1371/journal.pone.0000637

Reinhart, R. M. G., Mathalon, D. H., Roach, B. J., & Ford, J. M. (2011). Relationships between prestimulus gamma power and subsequent p300 and reaction time breakdown in schizophrenia. *International Journal of Psychophysiology*, *79*(1), 16–24. https://doi.org/10.1016/j.ijpsycho.2010.08.009

Ryan, D. B., Townsend, G., Gates, N. A., Colwell, K., & Sellers, E. W. (2017). Evaluating brain-computer interface performance using color in the p300 checkerboard speller. *Clinical Neurophysiology*, *128*(10), 2050–2057. https://doi.org/10.1016/j.clinph.2017.07.397

Salazar, V. (2020). *Theoretical and experimental study of p300 ERP in the context of brain-computer interfaces. part i: Study and analysis of functional connectivity methods.* (Master Thesis). Universidad Autónoma de Madrid. Escuela Politécnica Superior.

Satomi, K., Horai, T., Kinoshita, Y., & Wakazono, A. (1995). Hemispheric asymmetry of event-related potentials in a patient with callosal disconnection syndrome: A comparison of auditory, visual and somatosensory modalities. *Electroencephalography and Clinical Neurophysiology*, *94*(6), 440–449. https://doi.org/10.1016/0013-4694(94)00314-B

Schalk, G., McFarland, D., Hinterberger, T., Birbaumer, N., & Wolpaw, J. (2004). BCI2000: A general-purpose brain-computer interface (BCI) system. *IEEE Transactions on Biomedical Engineering*, *51*(6), 1034–1043. https://doi.org/10.1109/TBME.2004.827072

Schultz, D. H., & Cole, M. W. (2016). Higher intelligence is associated with less task-related brain network reconfiguration. *Journal of Neuroscience*, *36*(33), 8551–8561. https://doi.org/10.1523/JNEUROSCI.0358-16.2016

Schwent, V. L., Snyder, E., & Hillyard, S. A. (1976). Auditory evoked potentials during multichannel selective listening: Role of pitch and localization cues. *Journal of Experimental Psychology. Human Perception and Performance*, *2*(3), 313–325. https://doi.org/10.1037//0096-1523.2.3.313

Sellers, E. W., & Donchin, E. (2006). A p300-based brain–computer interface: Initial tests by ALS patients. *Clinical Neurophysiology*, *117*(3), 538–548. https://doi.org/10.1016/j.clinph.2005.06.027

Serby, H., Yom-Tov, E., & Inbar, G. (2005). An improved p300-based brain-computer interface. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, *13*(1), 89–98. https://doi.org/10.1109/TNSRE.2004.841878

Serrano-Marugán, I., Herrera, B., Romero, S., Nogales, R., Poch-Broto, J., Quintero, J., & Ortiz, T. (2014). [passive tactile stimulation and its clinical and neurophysiological repercussions (p300) in blind children with symptoms of attention deficit disorder]. *Revista De Neurologia*, *58 Suppl 1*, S25–30.

Sharbrough, F., Chatrian, G. E., Lesser, R., Lüders, H., Nuwer, M., & Picton, T. (1991). American electroencephalographic society guidelines for standard electrode position nomenclature. *Journal of clinical neurophysiology : official publication of the American Electroencephalographic Society*. https://doi.org/10.1097/00004691-199104000-00007

Sheng, Y., Liu, S., Wang, W., He, Y., Liu, X., Ke, Y., An, X., & Ming, D. (2018). A study on RSVP paradigm based on brain computer interface across subjects. *2018 9th International Conference on Awareness Science and Technology (iCAST)*, 42–46. https://doi.org/10.1109/ICAwST.2018.8517249

Sidorov, G., Gelbukh, A., Gómez-Adorno, H., & Pinto, D. (2014). Soft similarity and soft cosine measure: Similarity of features in vector space model [Number: 3]. *Computación y Sistemas*, *18*(3), 491–504–504. https://doi.org/10.13053/cys-18-3-2043

Siegler, R. S. (2007). Cognitive variability. *Developmental Science*, *10*(1), 104–109. https://doi.org/https://doi.org/10.1111/j.1467-7687.2007.00571.x

Stuss, D. T., Pogue, J., Buckle, L., & Bondar, J. (1994). Characterization of stability of performance in patients with traumatic brain injury: Variability and consistency on reaction time tests. *Neuropsychology*, *8*(3), 316–324. https://doi.org/10.1037/0894-4105.8.3.316

Tarvainen, M. P., Niskanen, J. .-P., Lipponen, J. A., Ranta-aho, P. O., & Karjalainen, P. A. (2009). Kubios HRV — a software for advanced heart rate variability analysis. In J. Vander Sloten, P. Verdonck, M. Nyssen, & J. Haueisen (Eds.), *4th european conference of the international federation for medical and biological engineering* (pp. 1022–1025). Springer. https://doi.org/10.1007/978-3-540-89208-3_243

Thomas, K. P., & Vinod, A. P. (2016). Biometric identification of persons using sample entropy features of EEG during rest state. *2016 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, 003487–003492. https://doi.org/10.1109/SMC.2016.7844773

Townsend, G., LaPallo, B. K., Boulay, C. B., Krusienski, D. J., Frye, G. E., Hauser, C. K., Schwartz, N. E., Vaughan, T. M., Wolpaw, J. R., & Sellers, E. W. (2010). A novel p300-based brain-computer interface stimulus presentation paradigm: Moving beyond rows and columns. *Clinical Neurophysiology: Official Journal of the International Federation of Clinical Neurophysiology*, *121*(7), 1109–1120. https://doi.org/10.1016/j.clinph.2010.01.030

Treder, M. S., & Blankertz, B. (2010). (c)overt attention and visual speller design in an ERP-based brain-computer interface. *Behavioral and brain functions: BBF*, *6*, 28. https://doi.org/10.1186/1744-9081-6-28

van Dinteren, R., Arns, M., Jongsma, M. L. A., & Kessels, R. P. C. (2014). P300 development across the lifespan: A systematic review and meta-analysis. *PLoS ONE*, *9*(2). https://doi.org/10.1371/journal.pone.0087347

Vaquero, E., Cardoso, M. J., Vazque, M., & Gomez, C. M. (2004). Gender differences in event-related potentials during visual-spatial attention. *International Journal of Neuroscience*, *114*(4), 541–557. https://doi.org/10.1080/00207450490422056

Verleger, R. (1997). On the utility of p3 latency as an index of mental chronometry. *Psychophysiology*, *34*(2), 131–156. https://doi.org/10.1111/j.1469-8986.1997.tb02125.x

Vidal, J. J. (1977). Real-time detection of brain events in EEG. *Proceedings of the IEEE*, *65*(5), 633–641. https://doi.org/10.1109/PROC.1977.10542

Wickens, C., Kramer, A., Vanasse, L., & Donchin, E. (1983). Performance of concurrent tasks: A psychophysiological analysis of the reciprocity of information-processing resources. *Science (New York, N.Y.)*, *221*(4615), 1080–1082. https://doi.org/10.1126/science.6879207

Wilcox, R. (2012). *Introduction to robust estimation and hypothesis testing*. Elsevier. https://doi.org/10.1016/C2010-0-67044-1

Wolpaw, J. R., Birbaumer, N., McFarland, D. J., Pfurtscheller, G., & Vaughan, T. M. (2002). Brain–computer interfaces for communication and control. *Clinical Neurophysiology*, *113*(6), 767–791. https://doi.org/10.1016/S1388-2457(02)00057-3

Xu, M., Liu, J., Chen, L., Qi, H., He, F., Zhou, P., Wan, B., & Ming, D. (2016). Incorporation of inter-subject information to improve the accuracy of subject-specific p300 classifiers. *International Journal of Neural Systems*, *26*(3), 1650010. https://doi.org/10.1142/S0129065716500106

Yamaguchi, S., & Knight, R. T. (1991). P300 generation by novel somatosensory stimuli. *Electroencephalography and Clinical Neurophysiology*, *78*(1), 50–55. https://doi.org/10.1016/0013-4694(91)90018-Y

Zhang, F. (2011). *Cross-validation and regression analysis in high-dimensional sparse linear models*. Stanford University.

Zhang, R., Yao, D., Valdés-Sosa, P. A., Li, F., Li, P., Zhang, T., Ma, T., Li, Y., & Xu, P. (2015). Efficient resting-state EEG network facilitates motor imagery performance. *Journal of Neural Engineering*, *12*(6), 066024. https://doi.org/10.1088/1741-2560/12/6/066024

Appendices

A

Methods

A.1 R2 calculation results.

This appendix contains the r2 calculation performed for each subject using 6 sessions for each in the general analysis and 2 sessions per day during three experimental days for each subject in the analysis per day for both helmets used.

A.1.1 Enobio Helmet

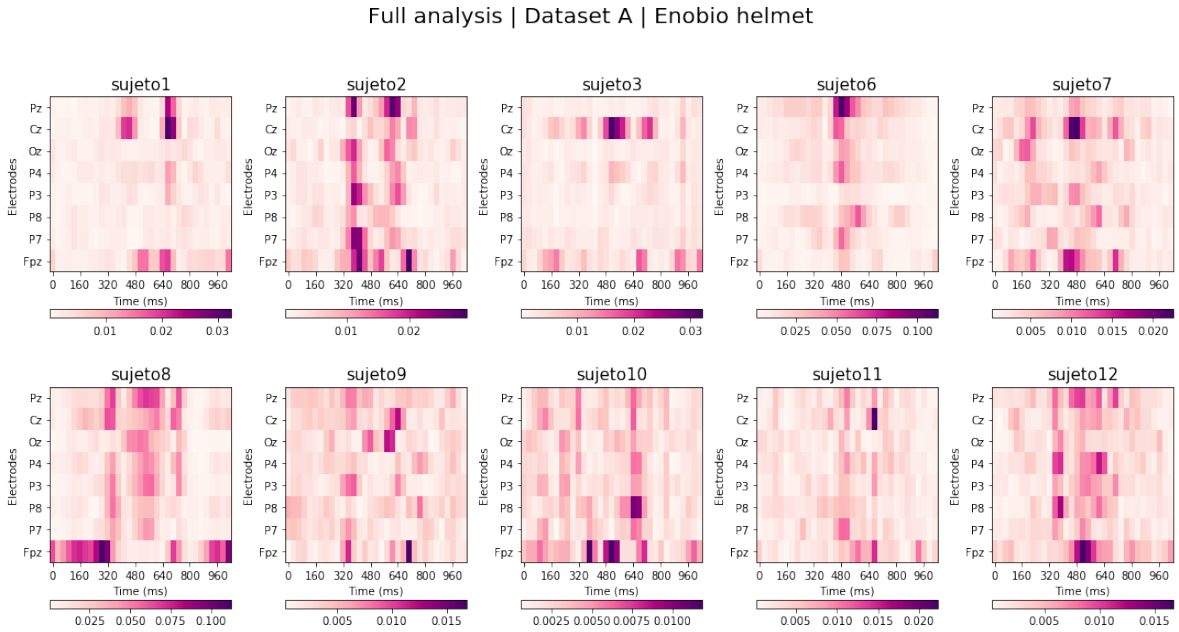


Figure A.1: R2 results of general analysis with Enobio helmet.

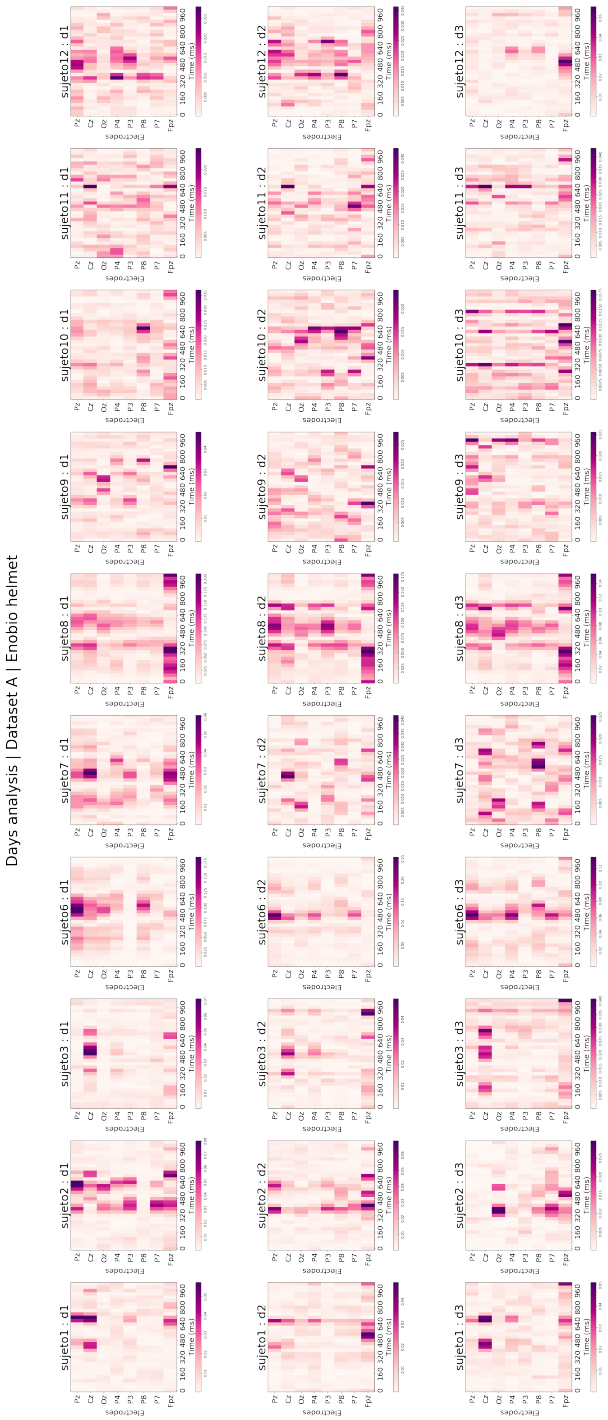


Figure A.2: R2 results of days analysis with Enobio helmet.

A.1.2 g-Tec Helmet

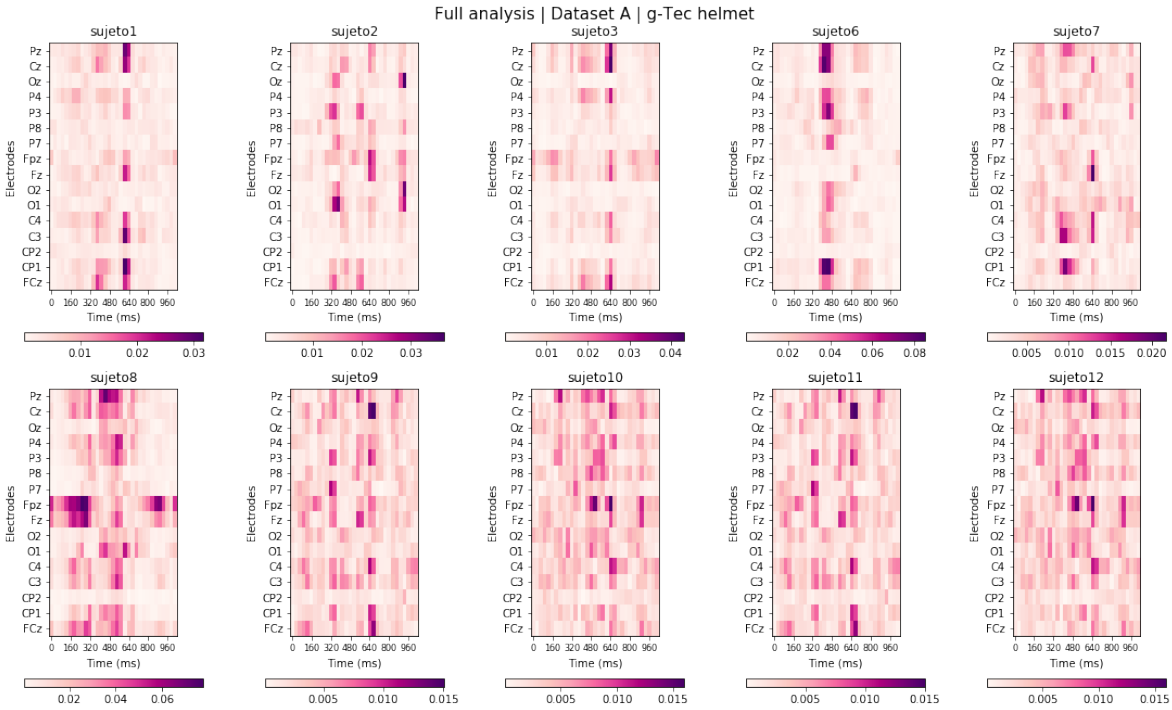


Figure A.3: R2 results of general analysis with g-Tec helmet.

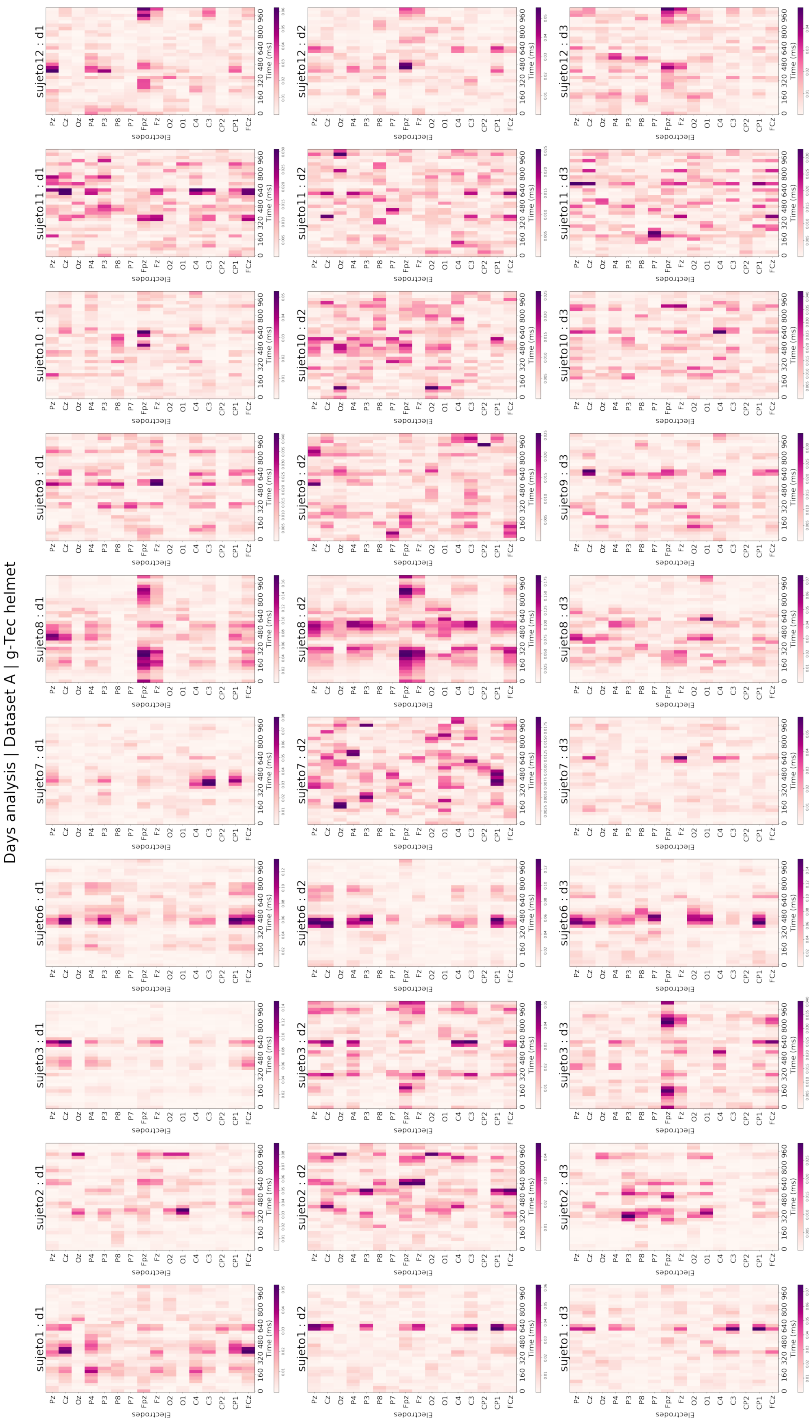


Figure A.4: R2 results of days analysis with g-Tec helmet.

B

Results

B.1 Inter-Subject Analysis

This appendix contains the accuracy results for the analysis by days in both helmets, varying the number of electrodes selected with BLDA-FS, with variations from 1 to 8. For each analyzed case, the number of electrodes that allow reaching the maximum accuracy value was defined as the optimal electrodes, this value is represented by the purple bar in the figures.

B.1.1 Analysis using cross-validation 2.

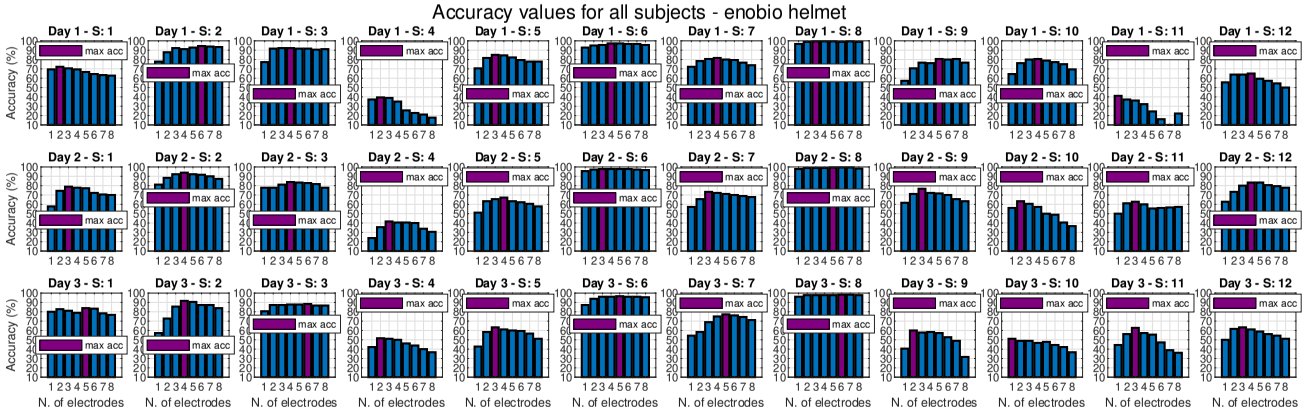


Figure B.1: Accuracy results of inter-subject variability analysis using BLDA electrodes for Enobio helmet and cross-validation 2.

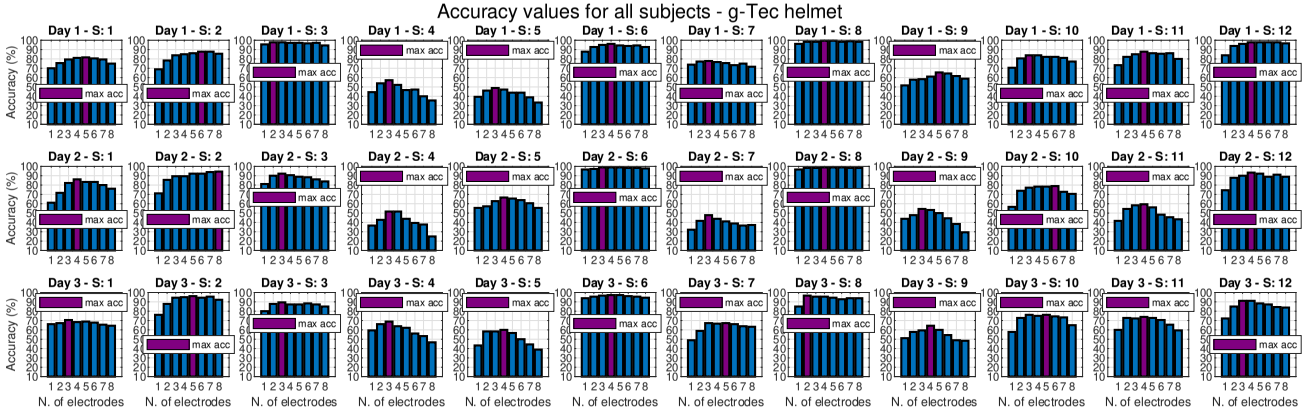


Figure B.2: Accuracy results of inter-subject variability analysis using BLDA electrodes for g-Tec helmet and cross-validation 2.

B.1.2 Analysis using cross-validation 12.

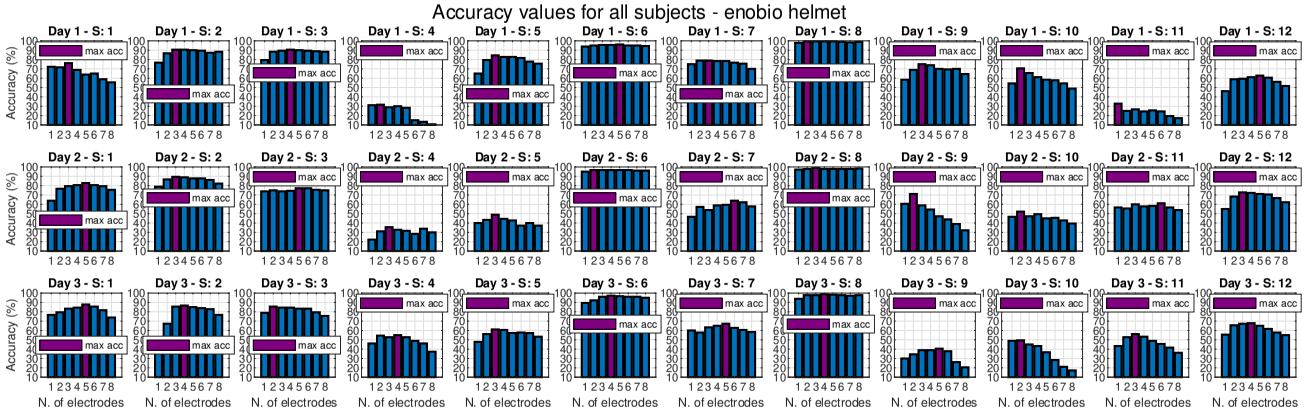


Figure B.3: Accuracy results of inter-subject variability analysis using BLDA electrodes for Enobio helmet cross-validation 12.

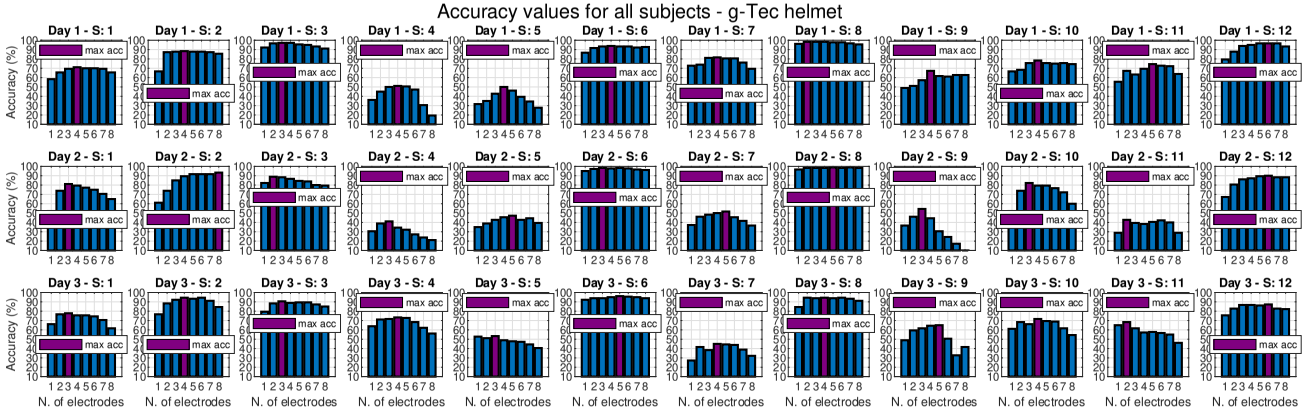


Figure B.4: Accuracy results of inter-subject variability analysis using BLDA electrodes for g-Tec helmet and cross-validation 2.

B.1.3 Analysis for cases.

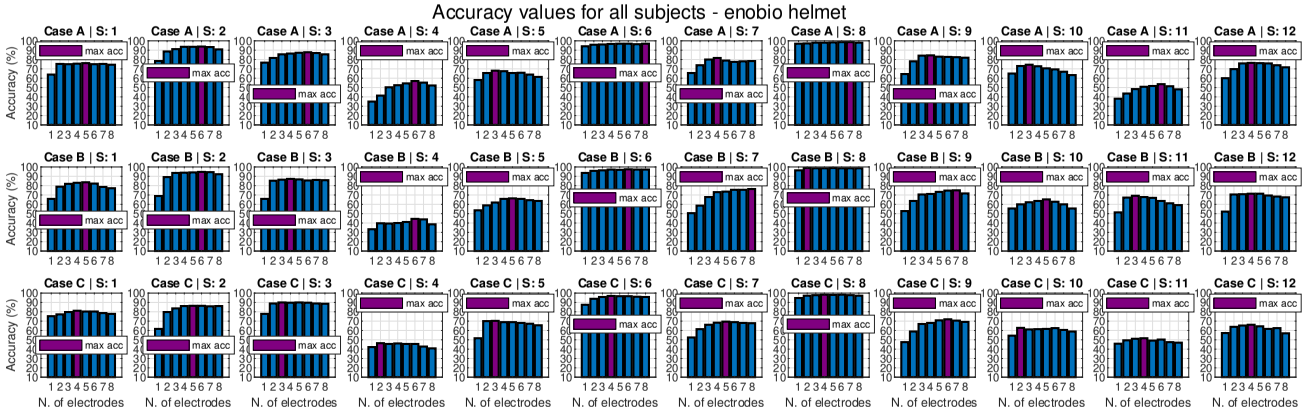


Figure B.5: Accuracy results of Cases A, B and C for Enobio helmet using BLDA electrodes.

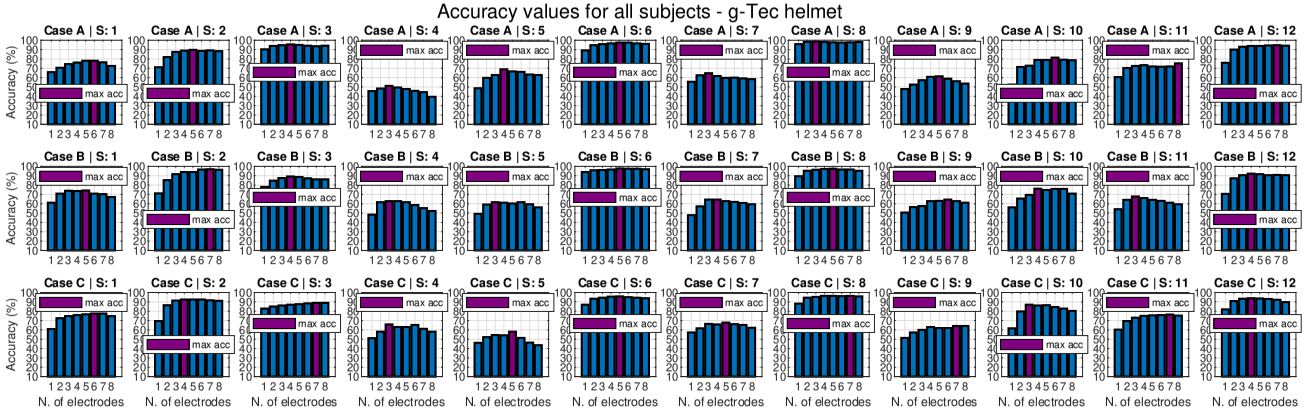


Figure B.6: Accuracy results of Cases A, B and C for g-Tec helmet using BLDA electrodes.